



Universidad Nacional de La Plata

Facultad de Informática

Tesis presentada para obtener el grado de
Doctor en Ciencias Informáticas

**EMPODERAMIENTO DE LA CONCIENCIA
SITUACIONAL EN OPERACIONES MILITARES
UTILIZANDO REALIDAD AUMENTADA**

Autor:

Mag. Alejandro Daniel Mitaritonna

Director:

Dra. María José Abásolo

Co-director:

Dr. Francisco Montero Simarro

La Plata, Buenos Aires, Diciembre de 2019

Dedicatoria

A mi hijo Germán fuente principal de amor, alegría, inspiración y motivación.

Agradecimientos

A mi directora de tesis por tanta paciencia y dedicación en el cumplimiento de mis objetivos.

A Paco, quién con su visión crítica y sus sugerencias me permitió concretar este trabajo de tesis.

Al personal civil y militar que me asistió en todo el proceso de desarrollo de esta tesis.

A mis compañeros de trabajo Carlos, Juan Carlos, Mariano, Darío, Diego y Cynthia que me apoyaron y ayudaron en realizar las pruebas finales.

Resumen

Durante las operaciones militares, los campos de batalla se convierten en zonas fracturadas donde el nivel de confusión, el ruido y la ambigüedad impactan en la manera de alcanzar los objetivos tácticos. La Conciencia Situacional (CS) se convierte en un reto ya que la percepción de la situación es inestable, lo que conduce a la comprensión degradada y a la incapacidad del soldado en proyectar los resultados apropiados. Para afrontar dicho reto diversos proyectos militares han centrado sus esfuerzos en diseñar un sistema digital integrado como soporte para la toma de decisiones del personal.

En esta tesis doctoral se propone un framework de Realidad Aumentada (RA) denominado RAIOM (Realidad Aumentada para la Identificación de Objetivos Militares) –entendiéndose por identificación de objetivos militares a todos los objetivos militares del enemigo circundantes a los soldados que representan una amenaza bélica- cuyo objetivo es mejorar la CS de los soldados en el campo de batalla mediante el uso de la RA. Para el desarrollo del framework se consideró como principal característica a la CS, y a la RA como tecnología de visualización e interpretación contextual. La aplicación de metodologías tales como SCRUM y Diseño Centrado en el Usuario (DCU) permitió identificar y analizar los requisitos funcionales y desarrollar los despliegues correspondientes. En particular, este trabajo presenta una recopilación actualizada de diferentes frameworks, menciona las características principales que debe poseer un framework de software de RA para uso militar, se detalla el uso de metodologías combinadas, se considera una arquitectura basada en capas para facilitar la integración tecnológica como también una arquitectura de sistemas distribuida donde la principal característica se centra en procesar las imágenes en un componente de hardware externo. Se realizan los despliegues utilizando el framework de RA y se obtienen los resultados provenientes de la evaluación de los despliegues que se centran en mejorar la CS. Finalmente, se detallan las conclusiones arribadas y se proponen líneas futuras de investigación.

Palabras claves: *realidad aumentada, conciencia situacional, framework, diseño centrado en el usuario, experiencia del usuario, SCRUM+DCU, SAGAT, SART, UMUX*

Índice de contenido

LISTA DE FIGURAS	XI
LISTA DE TABLAS.....	XV
TABLA DE ABREVIATURAS	XVII

I. ESTADO DE SITUACIÓN

CAPÍTULO 1	1
INTRODUCCIÓN.....	1
1.1 MOTIVACIÓN.....	1
1.2 OBJETIVOS.....	4
1.3 HIPÓTESIS	4
1.4 METODOLOGÍA DE INVESTIGACIÓN	5
1.5 ORGANIZACIÓN DEL INFORME.....	5
CAPÍTULO 2	9
CONCIENCIA SITUACIONAL EN PROYECTOS MILITARES.....	9
2.1 INTRODUCCIÓN	9
2.2 ¿QUÉ ES LA CS?	9
2.3 ANTECEDENTES DE PROYECTOS MILITARES.....	12
2.4 PROYECTOS MILITARES ACTUALES.....	16
2.5 RESUMEN DEL CAPÍTULO	19
CAPÍTULO 3	21
FUNDAMENTOS DE LA REALIDAD AUMENTADA.....	21
3.1 INTRODUCCIÓN	21
3.2 ¿QUÉ ES LA RA?.....	21
3.3 PARTES DE UNA APLICACIÓN DE RA	27
3.4 MÉTODOS DE ENTRADA DE DATOS EN UN SISTEMA DE RA.....	31
3.5 VISUALIZACIÓN	34
3.6 TECNOLOGÍAS PARA VISUALIZAR LA RA	39
3.7 TRACKING	41
3.7.1 Tracking mediante dispositivos físicos específicos	44

3.7.2 Tracking basado en visión	45
3.7.3 Tracking híbrido	50
3.8 RECONOCIMIENTO DE OBJETOS.....	51
3.9 ARQUITECTURA DE SOFTWARE PARA APLICACIONES DE RA	56
3.10 RESUMEN DEL CAPÍTULO.....	61
CAPÍTULO 4	63
CONCIENCIA SITUACIONAL Y REALIDAD AUMENTADA.....	63
4.1 INTRODUCCIÓN	63
4.2 LA CONCIENCIA SITUACIONAL Y LA REALIDAD AUMENTADA.....	63
4.3 REALIDAD AUMENTADA EN PROYECTOS MILITARES	65
4.3.1 Proyecto Eyekon.....	66
4.3.2 Proyecto BARS.....	67
4.3.3 Proyecto iARM	69
4.3.4 Proyecto ULTRA-Vis.....	70
4.4 COMPARATIVA DE TECNOLOGÍAS DIGITALES MILITARES QUE UTILIZAN RA	74
4.5 RESUMEN DEL CAPÍTULO	76
CAPÍTULO 5	79
ARQUITECTURA Y DISEÑO DE SOFTWARE	79
5.1 INTRODUCCIÓN	79
5.2 FRAMEWORK Y SU IMPORTANCIA	79
5.3 COMPARATIVA DE FRAMEWORKS DE RA MÁS UTILIZADOS	83
5.4 ARQUITECTURA DE SOFTWARE POR CAPAS Y COMPONENTES.....	85
5.5 AMBIENTE DE DESARROLLO	87
5.6 TÉCNICAS DE OPTIMIZACIÓN	90
5.6.1 Niveles de optimización.....	91
5.6.2 Optimización a nivel software.....	91
5.6.3 Optimización a nivel hardware	93
5.7 RESUMEN DEL CAPÍTULO	94
II. FRAMEWORK PROPUESTO	
CAPÍTULO 6	97
REQUERIMIENTOS INICIALES.....	97
6.1 INTRODUCCIÓN	97

6.2	ESQUEMA CONCEPTUAL DE OPERACIONES	97
6.3	IDENTIFICACIÓN, SELECCIÓN Y CLASIFICACIÓN DE REQUERIMIENTOS DE LA CS.....	100
6.4	RESUMEN DEL CAPÍTULO	109
CAPÍTULO 7		111
DISEÑO DE LAS ARQUITECTURAS.....		111
7.1	INTRODUCCIÓN	111
7.2	METODOLOGÍA DE DESARROLLO UTILIZADA	111
7.3	ARQUITECTURA DEL FRAMEWORK. MIDDLEWARE	115
7.3.1	Capa: Módulos	115
7.3.1.1	Módulo Input	118
7.3.1.2	Módulo Location.....	119
7.3.1.3	Módulo Mapping.....	120
7.3.1.4	Módulo Communication.....	121
7.3.1.5	Módulo Visualization	122
7.3.1.6	Módulo Vision.....	128
7.3.2	Capa: Librerías.....	134
7.4	ARQUITECTURA DEL SISTEMA. COMPONENTES DE HARDWARE UTILIZADOS	136
7.5	PROCESO DISTRIBUIDO.....	141
7.6	RESUMEN DEL CAPÍTULO	144
CAPÍTULO 8		147
DESPLIEGUE DE LA SOLUCIÓN		147
8.1	INTRODUCCIÓN	147
8.2	DISPOSITIVOS UTILIZADOS.....	148
8.3	DESPLIEGUES CONSIDERADOS.....	151
8.4	MÓDULO COMMUNICATION	154
8.5	MÓDULO VISION	155
8.6	INTEGRACIÓN DE TECNOLOGÍAS	157
8.7	RESUMEN DEL CAPÍTULO	159
CAPÍTULO 9		161
EVALUACIÓN		161
9.1	INTRODUCCIÓN	161
9.2	RESUMEN DE LA EVALUACIÓN.....	163
9.3	DESCRIPCIÓN DEL PRODUCTO SOFTWARE.....	163

9.4	OBJETIVOS DE LA EVALUACIÓN	164
9.5	MÉTODO UTILIZADO PARA LA EVALUACIÓN	164
9.5.1	Participantes	164
9.5.2	Tareas	165
9.5.3	Ejecución de las pruebas.....	168
9.5.4	Tecnología utilizada	169
9.5.5	Herramientas del Administrador de prueba	169
9.5.6	Diseño experimental	169
9.5.7	Procedimiento	171
9.6	PRESENTACIÓN Y DISCUSIÓN DE RESULTADOS	171
9.6.1	Prueba 1 realizada el 10 de noviembre de 2018.....	171
9.6.2	Prueba 2 realizada el 11 de noviembre de 2019.....	181
9.7	RESUMEN DEL CAPÍTULO.....	189
CAPÍTULO 10		191
CONCLUSIONES Y TRABAJOS FUTUROS.....		191
10.1	CONCLUSIONES.....	191
10.2	APORTES REALIZADOS.....	196
10.3	PUBLICACIONES.....	197
10.4	TRABAJOS FUTUROS	202
BIBLIOGRAFÍA.....		205

III. APÉNDICES

APÉNDICE A . CUADROS DE EVALUACIÓN

APÉNDICE B . TAREAS REALIZADAS POR CADA PARTICIPANTE

APÉNDICE C . EVALUACIONES DE PERFORMANCE DE DISPOSITIVOS MÓVILES

APÉNDICE D . DISEÑO DE PROTOTIPOS UTILIZANDO HISTORIAS DE USUARIOS

APÉNDICE E . VIDEOS DEMOSTRATIVOS

Lista de figuras

FIGURA 2.1 – MODELO DE LA CS EN LA TOMA DE DECISIÓN DINÁMICA	10
FIGURA 2.2 – MODELO DE LOS TRES NIVELES DE LA CS	12
FIGURA 2.3 – PROTOTIPO DE SOLDIER’S COMPUTER JUNTO CON SUS COMPONENTES ELECTRÓNICOS.....	13
FIGURA 2.4 – CONCEPTO DEL FUNCIONAMIENTO DEL SUBSISTEMA IHAS	15
FIGURA 2.5 – (A) PRIMER PROTOTIPO FUNCIONAL DEL HMD CREADO POR CRYE ASSOCIATED – (B) DISEÑO CONCEPTUAL DEL HMD CON VISIÓN TÉRMICA, CÁMARA DE VIDEO, SENSORES QBR Y UN HUD PARA MOSTRAR LOS DATOS E IMÁGENES.....	16
FIGURA 2.6 – PROGRAMAS SOLDADOS DEL FUTURO: (A) FIST – (B) COMFUT – (C) IdZ	19
FIGURA 3.1 – LA RA ABARCA UN AMPLIO ESPECTRO DE SISTEMAS. DENTRO DEL DOMINIO MAR.....	22
FIGURA 3.2 – CONTINUUM DE LA VIRTUALIDAD.....	23
FIGURA 3.3 – DIAGRAMA DE LAS PARTES DE UNA APLICACIÓN DE REALIDAD AUMENTADA	28
FIGURA 3.4 – MODELO DE REFERENCIA. COMPONENTES DE SOFTWARE REQUERIDOS EN UN SISTEMA DE RA UBICUA	29
FIGURA 3.5 – ARQUITECTURA DE REFERENCIA GENÉRICA DE SISTEMAS DE RA	30
FIGURA 3.6 – SISTEMA DE RECONOCIMIENTO GESTUAL	33
FIGURA 3.7 – SISTEMAS DE COORDENADAS DE UN SISTEMA DE RA.....	36
FIGURA 3.8 – DIAGRAMA CONCEPTUAL DE LA TECNOLOGÍA <i>OPTICAL SEE-THROUGH</i>	40
FIGURA 3.9 – DIAGRAMA CONCEPTUAL DE LA TECNOLOGÍA <i>VIDEO SEE-THROUGH</i>	40
FIGURA 3.10 – DIFERENTES TIPOS DE MARCADORES: A) TEMPLATE; B) ID-MARKER; C) DATAMATRIX Y D) QR.....	46
FIGURA 3.11 – MARCO DE REFERENCIA USADO PARA ESTIMACIÓN DE LA POSE	50
FIGURA 3.12 – DIFERENTES NIVELES DE PROCESAMIENTO DE LAS TAREAS DE RA. A) LAS TAREAS SE EJECUTAN EN EL CLIENTE; B) EN EL SERVER SE EJECUTA LA TAREA DE TRACKING; C) EN EL SERVIDOR SE EJECUTAN LAS TAREAS DE TRACKING Y LA LÓGICA DE LA APLICACIÓN Y D) TODAS LAS TAREAS SE REALIZAN EN EL SERVIDOR	57
FIGURA 3.13 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON RECONOCIMIENTO LOCAL	58
FIGURA 3.14 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON RECONOCIMIENTO Y TRACKING LOCAL	59
FIGURA 3.15 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON RECONOCIMIENTO REMOTO	59
FIGURA 3.16 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON REGISTRACIÓN LOCAL, RECONOCIMIENTO REMOTO Y TRACKING.....	60
FIGURA 3.17 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON TRACKING Y REGISTRACIÓN LOCAL Y RECONOCIMIENTO REMOTO	60
FIGURA 3.18 – REPRESENTACIÓN DE LA ARQUITECTURA DE UN SISTEMA DE RA PROCESANDO LA INFORMACIÓN CON TIEMPO REAL, REGISTRACIÓN REMOTA, DETECCIÓN REMOTA Y PRESENTACIÓN LOCAL.....	61
FIGURA 4.1 - DISEÑO CONCEPTUAL DEL PROYECTO IARM	69

FIGURA 4.2 – COMPONENTES DEL PROTOTIPO ULTRA-Vis.....	71
FIGURA 4.3 – INTERFAZ DE USUARIO DEL PROTOTIPO ULTRA-Vis	72
FIGURA 4.4 – INFORMACIÓN PROVISTA Y COMPONENTES DE HARDWARE DEL PROYECTO ACR4	73
FIGURA 5.1 – ARQUITECTURA DE ANDROID	88
FIGURA 5.2 – SECUENCIA DE EJECUCIÓN DE UNA APLICACIÓN ANDROID	89
FIGURA 5.3 – ARQUITECTURA COMÚN DE UN PROCESADOR CON MULTI-CORE, GPU, DSP, ETC., EN UN ÚNICO CHIP	93
FIGURA 6.1 – FLUJO DE INFORMACIÓN ENTRE BLOQUES	98
FIGURA 6.2 – ESQUEMA DE LA UBICACIÓN DE LOS INTEGRANTES/USUARIOS PRINCIPALES	99
FIGURA 6.3 – GRÁFICO DE LAS PREGUNTAS ESPECÍFICAS DE GUTWIN ET AL. [70], EN CADA NIVEL DE LA CS DEL MODELO DE ENDSLEY	104
FIGURA 6.4 – COMBINACIÓN DE PROPUESTAS PARA LA DEFINICIÓN DE UN NUEVO MODELO PARA LA IDENTIFICACIÓN, SELECCIÓN Y CLASIFICACIÓN DE LOS REQUISITOS DE LA CS.....	107
FIGURA 6.5 – EJEMPLO DE USO DEL MODELO TRIDIMENSIONAL PROPUESTO PARA LA IDENTIFICACIÓN, SELECCIÓN Y CLASIFICACIÓN DE REQUISITOS DE LA CS	108
FIGURA 7.1 – SCRUM + DCU	113
FIGURA 7.2 – DISEÑO BASADO EN COMPONENTES DE LA ARQUITECTURA DEL FRAMEWORK	117
FIGURA 7.3 – DIAGRAMA DE CLASES DEL MÓDULO INPUT.....	119
FIGURA 7.4 – DIAGRAMA DE CLASES DEL MÓDULO LOCATION.....	120
FIGURA 7.5 – DIAGRAMA DE CLASES DEL MÓDULO MAPPING.....	121
FIGURA 7.6 – DIAGRAMA DE CLASES DEL MÓDULO COMMUNICATION	122
FIGURA 7.7 – FILTRADO DE INFORMACIÓN. CIERTO CONJUNTO DE INFORMACIÓN ES PRIORITARIO. SE EXCLUYE EL RESTO DE INFORMACIÓN.....	123
FIGURA 7.8 – SIMBOLOGÍA DEL RADAR. ESTRUCTURA DE LA VISUALIZACIÓN DEL RADAR.....	124
FIGURA 7.9 – MAPA 360 GRADOS DEL PROTOTIPO DESARROLLADO UTILIZANDO EL FRAMEWORK.....	125
FIGURA 7.10 – SIMBOLOGÍA GESTUAL	126
FIGURA 7.11 – DIAGRAMA DE CLASES DEL MÓDULO VISUALIZATION	126
FIGURA 7.12 – DIAGRAMA DE CLASES DE LOS MÓDULOS DEL FRAMEWORK	127
FIGURA 7.13 – PRUEBAS INICIALES DE RECONOCIMIENTO GESTUAL Y VISUALIZACIÓN DE ICONOS	129
FIGURA 7.14 – EXPLICACIÓN GRÁFICA DE BAG OF WORDS: (i) IMÁGENES REPRESENTANDO UNA PALABRA VISUAL (ii) CARACTERÍSTICAS (iii) DESCRIPTORES UBICADOS EN EL ESPACIO (iv) AGRUPACIÓN EN CLUSTERS (v) IMAGEN A CLASIFICAR (vi) CARACTERÍSTICAS (vii) CORRESPONDENCIAS CON LOS DESCRIPTORES CONOCIDOS (viii) HISTOGRAMA DE CORRESPONDENCIAS	131
FIGURA 7.15 –PRUEBA DE IDENTIFICACIÓN DEL EDIFICIO DE ELECTRÓNICA DEL CITEDEF UTILIZANDO ORB.....	133
FIGURA 7.16 – DIAGRAMA DE CLASES DEL MÓDULO VISION	134
FIGURA 7.17 – ARQUITECTURA PROPUESTA DEL SISTEMA UTILIZANDO LA VISTA DE COMPONENTES Y CONECTORES	136
FIGURA 7.18 – ARQUITECTURA DEL SISTEMA DE RA PROPUESTO	137
FIGURA 7.19 – DISEÑO DE LAS CAPACIDADES DEL FRAMEWORK PROPUESTO.....	138
FIGURA 7.20 – EJEMPLO DE LA ARQUITECTURA PROPUESTA UTILIZANDO GAFAS DE RA (1) Y ODROID XU3 (2)	139

FIGURA 7.21 – ODROID XU3. SOLUCIÓN HETEROGENEOUS MULTI-PROCESSING (HMP) SOLUTION	140
FIGURA 7.22 – ODROID XU3	141
FIGURA 7.23 – ODROID XU3 IMPLEMENTADO EN EL LABORATORIO DEL CITEDEF.....	141
FIGURA 7.24 – DIAGRAMA DE DESPLIEGUE	142
FIGURA 7.25 – EXTRAYENDO CARACTERÍSTICAS DE UNA IMAGEN MEDIANTE PROCESAMIENTO DISTRIBUIDO UTILIZANDO UN ODROID XU3 Y UNA PC QUE SIMULA A UN DISPOSITIVO MÓVIL	143
FIGURA 8.1 – DISPOSITIVOS MÓVILES UTILIZADOS PARA EN LOS DESPLIEGUES	151
FIGURA 8.2 – <i>GLASS CARDBOX</i> ACOPLADO A LOS SMARTPHONE UTILIZADO EN LOS DESPLIEGUES	152
FIGURA 8.3 – OPCIONES DE ARQUITECTURAS DE SISTEMAS PROPUESTOS.....	153
FIGURA 8.4 – MÓDULO CLIENTE Y MÓDULO SERVIDOR.....	154
FIGURA 8.5 – PRUEBA DE RECONOCIMIENTO GESTUAL.....	155
FIGURA 8.6 – PRUEBA DE EXTRACCIÓN DE CARACTERÍSTICAS UTILIZANDO ORB. RECONOCIMIENTO Y CLASIFICACIÓN DE OBJETOS EDIFICIOS	156
FIGURA 8.7 – VISUALIZACIÓN DE ICONOGRAFÍA Y RADAR DE 360 GRADOS CON INFORMACIÓN CONTEXTUAL IMPRESA	156
FIGURA 8.8 – ACTIVACIÓN DE UN MAPA POR MEDIO DE RECONOCIMIENTO GESTUAL.....	157
FIGURA 8.9 – QR PARA ACCEDER AL VÍDEO EN EL CUAL SE MUESTRA LA INTEGRACIÓN DE LAS FUNCIONALIDADES PRINCIPALES DEL FRAMEWORK.....	159
FIGURA 9.1 – TAXONOMÍA DE BLOOM	170
FIGURA 9.2 – GRAFICO DE EVALUACIÓN DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DEL GRUPO A DE PARTICIPANTES EVALUADOS.....	174
FIGURA 9.3 – GRAFICO DE EVALUACIÓN DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DEL GRUPO B DE PARTICIPANTES EVALUADOS.....	175
FIGURA 9.4 – GRAFICO DE EVALUACIÓN DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DE TODOS LOS PARTICIPANTES EVALUADOS (GRUPO A Y GRUPO B).....	176
FIGURA 9.5 – GRAFICO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX CORRESPONDIENTE A LOS PARTICIPANTES DEL GRUPO A.....	178
FIGURA 9.6 – GRAFICO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX CORRESPONDIENTE A LOS PARTICIPANTES DEL GRUPO B.....	179
FIGURA 9.7 – GRAFICO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX CORRESPONDIENTE A TODOS LOS PARTICIPANTES (GRUPO A Y GRUPO B).....	181
FIGURA 9.8 – GRAFICO DE EVALUACIÓN ACTUALIZADO DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DEL GRUPO A DE PARTICIPANTES EVALUADOS.....	182
FIGURA 9.9 – GRAFICO DE EVALUACIÓN ACTUALIZADO DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DEL GRUPO B DE PARTICIPANTES EVALUADOS.....	183
FIGURA 9.10 – GRAFICO DE EVALUACIÓN ACTUALIZADO DE LA CS POR NIVELES SEGÚN ENDSLEY: PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN DE TODOS LOS PARTICIPANTES EVALUADOS (GRUPO A Y GRUPO B).....	184
FIGURA 9.11 – GRAFICO ACTUALIZADO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX CORRESPONDIENTE A LOS PARTICIPANTES DEL GRUPO A.....	185

FIGURA 9.12 – GRÁFICO ACTUALIZADO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX CORRESPONDIENTE A LOS PARTICIPANTES DEL GRUPO B.....	187
FIGURA 10.14 – GRÁFICO DE CUMPLIMIENTO ACTUALIZADO DE LA CS DE TODOS LOS PARTICIPANTES (GRUPO A Y GRUPO B).....	188
FIGURA 10.1 – PROTOTIPO DE HMD DISEÑADO Y DESARROLLADO EN EL CITEDEF.....	203
FIGURA 10.2 – GRÁFICO COMPARATIVO DE PERFORMANCE ENTRE LOS DISPOSITIVOS MEDIDOS EN SEGUNDOS EN EL EVENTO <i>ONPREVIEWFRAME</i>	220
FIGURA 10.3 – GRÁFICO COMPARATIVO DE PERFORMANCE DEL PROCESAMIENTO DISTRIBUIDO EVALUADO EN UN DISPOSITIVO HUAWEI G8 Y EN UNA Gafa EPSON MOVERIO BT 300	223
FIGURA 10.4 – TIEMPO PROMEDIO DE EJECUCIÓN DEL ALGORITMO ORB EN LOS DISPOSITIVOS EN MILISEGUNDOS.....	224
FIGURA 10.5 – TIEMPO DE EJECUCIÓN DEL ALGORITMO ORB POR CADA UNA DE LAS 20 PRUEBAS REALIZADAS EN LOS DISPOSITIVOS MEDIDO EN MILISEGUNDOS	225
FIGURA 10.6 – RADAR RAIOM DE 360 GRADOS. 3D-SAM, DISEÑO E IMPLEMENTACIÓN DE PROTOTIPOS.....	228
FIGURA 10.7 – RECONOCIMIENTO DE PUNTOS DE REFERENCIA RAIOM. 3D-SAM, DISEÑO E IMPLEMENTACIÓN DE PROTOTIPOS	229
FIGURA 10.8 – MAPA INTERACTIVO DE RAIOM. 3D-SAM, DISEÑO E IMPLEMENTACIÓN DE PROTOTIPOS.....	230

Lista de tablas

TABLA 4.1 ELEMENTOS IMPRESCINDIBLES PARA MANTENER UN ALTO NIVEL DE CS	75
TABLA 4.2 TECNOLOGÍAS UTILIZADAS COMO SOPORTE A LA CS	76
TABLA 5.1 PARÁMETROS FUNCIONALES TECNOLÓGICOS.....	84
TABLA 5.2 PARÁMETROS NO FUNCIONALES.....	85
TABLA 6.1 ELEMENTOS DE LA CONCIENCIA DEL ESPACIO DE TRABAJO RELACIONADOS CON EL PRESENTE.....	101
TABLA 6.2 ELEMENTOS DE LA CONCIENCIA DEL ESPACIO DE TRABAJO RELACIONADOS CON EL PASADO.....	102
TABLA 6.3 TABLA DE DEFINICIÓN DE CU POR NIVELES DEL MODELO DE ENDSLEY.....	104
TABLA 7.1 VENTAJAS EN LA COMBINACIÓN DE LAS METODOLOGÍAS SCRUM Y DCU (SCRUM+DCU)	114
TABLA 7.2 RELACIÓN EXPLÍCITA ENTRE LA CARACTERIZACIÓN DE LA CS Y LOS MÓDULOS PROPUESTOS.....	118
TABLA 8.1 DISPOSITIVOS MÓVILES. CARACTERÍSTICAS TÉCNICAS.	149
TABLA 8.2 DESPLIEGUES UTILIZADOS PARA VALIDAR LA SOLUCIÓN PROPUESTA.	152
TABLA 9.1 PERFIL DE LOS PARTICIPANTES.	165
TABLA 9.2 TAREAS QUE DEBÍAN REALIZAR LOS PARTICIPANTES.	166
TABLA 9.3 PREGUNTAS ELABORADAS PARA COMPROBAR EL GRADO DE EFECTIVIDAD Y EFICIENCIA CON EL QUE LOS PARTICIPANTES PERCIBÍAN, COMPRENDÍAN O PROYECTABAN EN UN CONTEXTO DADO.	167
TABLA 9.4 COMANDOS DE VOZ Y ACCIÓN.	168
TABLA 9.5 RESULTADOS DE LA EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRESIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY.....	172
TABLA 9.6 RESULTADOS DE LA EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX ...	172
TABLA 9.7 CUADRO DE EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRESIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY. GRUPO A.....	173
TABLA 9.8 CUADRO DE EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRESIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY. GRUPO B.....	175
TABLA 9.9 CUADRO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX. GRUPO A	177
TABLA 9.10 CUADRO DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX. GRUPO B	178
TABLA 9.11 PERFIL DE LOS 4 PARTICIPANTES ADICIONALES AL EXPERIMENTO INICIAL.....	181
TABLA 9.12 CUADRO ACTUALIZADO (P4) DE EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRESIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY. GRUPO A.	182
TABLA 9.13 CUADRO ACTUALIZADO (P4, P5 Y P6) DE EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRESIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY. GRUPO B.	183
TABLA 9.14 CUADRO ACTUALIZADO (P4) DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX. GRUPO A.....	184
TABLA 9.15 CUADRO ACTUALIZADO (P4, P5, P6) DE EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX. GRUPO B	186

TABLA 9.16 RESULTADOS GLOBALES ACTUALIZADOS DE LA EVALUACIÓN DE LA CS POR NIVELES (PERCEPCIÓN, COMPRENSIÓN Y PROYECCIÓN) PROPUESTO POR ENDSLEY	187
TABLA 9.17 RESULTADOS GLOBALES ACTUALIZADOS DE LA EVALUACIÓN DE LA EXPERIENCIA DEL USUARIO BASADO EN EL CUESTIONARIO UMUX	188
TABLA 10.1 ARQUITECTURA DE LOS DISPOSITIVOS MÓVILES EVALUADOS.....	219
TABLA 10.2 COMPARACIÓN DE PERFORMANCE CORRESPONDIENTE AL TIEMPO DE CAPTURA DE VIDEO Y VISUALIZACIÓN EN LOS DISPOSITIVOS MÓVILES UTILIZADOS EN LOS DESPLIEGUES MEDIDO EN MILLISEGUNDOS	219
TABLA 10.3 CARACTERÍSTICAS PRINCIPALES DE LOS DISPOSITIVOS DE HARDWARE UTILIZADOS EN LA EVALUACIÓN	222
TABLA 10.4 COMPARACIÓN DE PERFORMANCE CORRESPONDIENTE AL PROCESAMIENTO DISTRIBUIDO MEDIDO EN MILLISEGUNDOS.....	222
TABLA 10.5 COMPARACIÓN DE PERFORMANCE PROMEDIO CORRESPONDIENTE AL PROCESAMIENTO DEL ALGORITMO ORB EN MILLISEGUNDOS	224

Tabla de abreviaturas

AAM	Active Appearance Model
API	Application Programming Interface
ARA	Applied Research Associates
ARM	Advanced RISC Machines
ASCII	American Standard Code for Information Interchange
BARS	Battlefield Augmented Reality System
BoW	Bag Of Words
BRIEF	Binary Robust Independent Elementary Features
BRISK	Binary Robust Invariant Scalable Keypoints
C4I2	Comando, Control, Comunicaciones, Computación, Información, Inteligencia
C4ISR	Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance
CAR	Contact-lens augmented reality
CC	Comando y Control
CECOM	Communications Electronics Command
CII	Centro Integrador
CITEDEF	Instituto de Investigaciones Científicas y Técnicas para la Defensa
COMFUT	Combatiente Futuro
COP	Common Operating Picture
COTS	Commercial Off-The-Shelf
CPU	Central Processing Unit
CS	Conciencia Situacional
CU	Caso de Uso
DARPA	Defense Advanced Research Projects Agency
DCU	Diseño Centrado en el Usuario
DoG	Difference of Gaussians
DoH	Determinant of Hessian
DSR	Design Science Research
DWARF	Distributed Wearable Augmented Reality
EE.AA.	Ejército Argentino
FAST	Features from Accelerated Segment Test
FELIN	Fantassin à Équipements et Liaisons Intégrés
FF.AA.	Fuerzas Armadas Argentinas
FIST	Future Integrated Soldier Technology
FREAK	Fast Retina Keypoint
GIS	Geographic Information System
GLOH	Gradient Location and Orientation Histogram
GPGPU	General-Purpose Computing on Graphics Processing Units
GPS	Global Positioning System
GPU	Graphics Processing Units
HCI	Human-Computer Interaction
HMAR	Handheld Mobile Augmented Reality

HMD	Head-Mounted Display
HMP	Heterogeneous Multi-Processing
HSV	Hue, Saturation, Value
HUD	Head-Up Displays
iARM	Intelligent Augmented Reality Model
IdZ	Infanterist der Zukunft
IHAS	Integrated Helmet Assembly Subsystem
IMU	Inertial Measurement Unit
IoC	Inversion of Control
JCa	Puesto de Mando
JIT	Just In Time
JNI	Java Native Interface
JPat	Jefe de Patrulla
LoG	Laplacian of Gaussian
LW	Land Warrior
MAR	Mobile Augmented Reality
MSER	Maximally Stable Extremal Regions
MVC	Model-View-Controller
MVD	Máquina Virtual Dalvik
MWS	Modular Weapon System
NDK	Native Development Kit
NRL	Naval Research Laboratory
OHA	Open Handset Alliance
ORB	Oriented FAST and Rotated BRIEF
Pat	Patrulla
PCA	Principal Component Analysis
PCBR	Principal Curvature-Based Region detector
POO	Programación Orientada a Objetos
PTAM	Parallel Tracking And Mapping
QBR	Químico, Biológico y Radioactivo
QR	Quick Response code
RA	Realidad Aumentada
RAIOM	Realidad Aumentada para la Identificación de Objetivos Militares
RFID	Radio Frequency Identification
RGP	Recognized Ground Picture
RPIDFA	Régimen del Personal de Investigación y Desarrollo de las FF.AA.
RTP	Real-time Transport Protocol
RV	Realidad Virtual
SAGAT	Situation Awareness Global Assessment Technique
SART	Situation Awareness Rating Technique
SPAM	Situation Present Awareness Method
SDK	Software Development Kit
SfM	Structure from Motion
SIFT	Scale-invariant feature transform
SIMD	Single Instruction Multiple Data
SLAM	Simultaneous Localization And Mapping

SMS	Short Message Service
SURF	Speeded Up Robust Features
UAV	Unmanned Aerial Vehicle
ULTRA-Vis	Urban Leader Tactical Response, Awareness and Visualization
UML	Unified Modeling Language
UNLP	Universidad Nacional de La Plata
URL	Uniform Resource Locator
WA	Workspace Awareness
XMPP	Extensible Messaging and Presence Protocol

I. Estado de Situación

Capítulo 1

Introducción

En este capítulo se da inicio a la primera Parte del trabajo tesis que se centra en detallar el Estado de Situación. Se explica la motivación, objetivos, hipótesis y la metodología de investigación adoptada para organizar el trabajo de tesis y por último la organización del informe, las partes y las secciones que conforman al presente trabajo.

1.1 Motivación

El fin de la Guerra Fría y el creciente número de conflictos de baja intensidad y pequeña escala aumenta la importancia de las tropas desmontadas. La infantería tendrá cada vez más un papel importante en los conflictos futuros. La necesidad de contar con una infantería bien equipada y entrenada será vital en el desempeño de las misiones en tierra. El avance en la tecnología otorgará a los potenciales enemigos contar con equipamiento de última generación permitiendo ser más letales y estar más protegidos. Las operaciones probablemente ocurrirán en lugares desconocidos. Los soldados tendrán que ejecutar cada vez más tareas. Es imprescindible contar con nuevos proyectos tecnológicos que apoyen las actividades de la infantería en misiones militares.

El equipamiento del futuro soldado debe consistir en un conjunto integrado de componentes electrónicos modulares para mejorar, de manera equilibrada, sus habilidades de comunicación, observación, letalidad, movilidad y protección. La interacción del soldado con el medio ambiente también significa que debe ser capaz de procesar la información contextual sin que afecte su capacidad cognitiva. Esto dio lugar a una variedad de programas de Sistemas de "Future Soldier" [1] lanzados por diferentes ejércitos. Estos programas tienen un alto grado de diferencia de sofisticación tecnológica y desarrollo. En general, todos estos programas tienen como objetivo mejorar sustancialmente la eficacia de combate y supervivencia del soldado. Una regla de los programas de modernización de los soldados es

que el soldado no opera solo, sino como parte de un equipo para llevar a cabo una determinada tarea en un entorno determinado. Adicionalmente en [2] se menciona que el fratricidio sigue siendo una amenaza muy real en los campos de batalla actuales. Para hacer frente a esto, los militares han puesto mucho esfuerzo en el desarrollo de tecnologías de identificación de combate para mejorar la capacidad de los soldados en identificar al enemigo con precisión. Debido a la naturaleza asimétrica de los conflictos recientes, estos sistemas utilizan cada vez más los dispositivos móviles sensibles al contexto para mejorar esta capacidad.

Por otro lado como se detalla en [3], muchas de las operaciones militares se desarrollan en entornos desconocidos. Estos complejos campos de batalla son muy exigentes e introducen desafíos para el combatiente, entre ellos el hecho de tener una CS tridimensional amplia del terreno es vital para que la operación sea un éxito, minimizando los efectos colaterales. Un reto importante en el desarrollo de herramientas eficaces de apoyo de decisiones radica en la forma en que se presenta la información, siendo uno de los principales objetivos reducir la carga cognitiva en momentos de estrés para ello las soluciones tecnológicas apoyadas en mejorar la CS permiten a los soldados hacer un uso efectivo de la información variada en un contexto de batalla. La nueva tecnología ofrece métodos innovadores de obtener información contextual y representarla visualmente de una manera natural y no invasiva sin afectar el proceso cognitivo del combatiente.

Como se menciona en [2], la CS es una representación mental y comprensión de los objetos, eventos, gente, estados de los sistemas, interacciones, condiciones ambientales y cualquier otro tipo de factores de una situación específica que puedan afectar al desarrollo de las tareas humanas, bien sean complejas o dinámicas. Logrando mantener la conciencia situacional se potencia la adquisición, la representación, la interpretación y la utilización de cualquier información relevante con el objeto de poner sentido a los eventos que ocurren, pudiéndose anticipar a los acontecimientos futuros, dando la capacidad de poder tomar decisiones inteligentes y de poder mantener el control. La CS es crucial para la mayoría de las operaciones militares, pero tiene diferentes requisitos en los diferentes niveles de las operaciones militares y las diferentes funciones.

Por otro lado, las aplicaciones de RA son cada vez más populares debido a los notables progresos en los dispositivos de computación móviles, como celulares inteligentes,

asistentes digitales personales y computadoras portátiles ligeras [4]. La posibilidad de acceso a Internet en todas partes, incluso desde los dispositivos móviles, alimenta el desarrollo con la disponibilidad de los dispositivos electrónicos que combinan las tecnologías necesarias en una sola unidad integrada. Generalmente las aplicaciones de RA se ejecutan en un ambiente externo, y el tracking del usuario se basa en utilizar sistema de posicionamiento global (GPS), brújula electrónica y/o acelerómetros, así como también análisis del vídeo de la escena real capturada en tiempo real. Asimismo, puede requerir el acceso directo a información georeferenciada. La aplicación también debe permitir visualizar la realidad con la superposición de objetos 3D generados. El punto principal dentro del desarrollo de la RA es un sistema de seguimiento de movimiento o *Tracking System*. Desde el principio hasta ahora la RA se apoya en “Marcadores” o un arreglo de Marcadores dentro del campo de visión de las cámaras para que el dispositivo tenga un punto de referencia sobre el cual superponer las imágenes. Recién en los últimos años el desarrollo de RA sin marcadores (en inglés, markerless) está madurando, pues la gran capacidad computacional de los nuevos dispositivos tecnológicos permite reconocer formas geométricas, pudiendo el sistema de seguimiento captar un objeto y superponer información o una imagen [5].

De acuerdo a lo que mencionan Livingston et al [6], en lo que respecta al ámbito militar, recientemente se está utilizando la RA como soporte para mejorar la CS en la toma de decisiones en particular en operaciones militares. Muchas de las operaciones militares se desarrollan en entornos desconocidos. Estos complejos campos de batalla en 3D son muy exigentes e introducen muchos desafíos para el combatiente. Estos incluyen visibilidad limitada, falta de familiaridad con el medio ambiente, amenazas de francotirador, ocultamiento de fuerzas enemigas, mala comunicación y un problema de la localización e identificación de los enemigos y de las fuerzas aliadas. Para ello tener una CS amplia del terreno es vital para que la operación sea un éxito minimizando los efectos colaterales. De acuerdo a la investigación previa efectuada, el único centro de investigación en América que está llevando adelante proyectos de este tipo es el DARPA pues considera a la RA como una tecnología que ayudará a realizar con éxito las operaciones militares [3].

En este contexto se evidencia la falta, en el ámbito de la Defensa, en Argentina, de especialistas que puedan aportar los conocimientos necesarios para el desarrollo de un software que implemente la RA para ayudar a mejorar el desempeño militar en operaciones

militares. Por esto, la motivación principal de la tesis es contribuir a desarrollar un framework de software que utilice la RA a fin de desarrollar aplicaciones que se utilicen en el ámbito militar, considerando requisitos importantes en este contexto, como son la CS, las facilidades de uso y la interoperabilidad y portabilidad de los artefactos desarrollados.

1.2 Objetivos

El objetivo de la tesis es aportar los conocimientos necesarios sobre CS en el ámbito de la Defensa para que sean implementados en soluciones concretas como apoyo a las actividades militares. Para lograr lo antes mencionado se propone desarrollar un framework basado en RA y centrado en el desarrollo de aplicaciones móviles con una arquitectura distribuida e integrador de tecnologías.

Por lo tanto, esto último trae aparejado los siguientes objetivos específicos:

- Explorar y analizar la bibliografía existente sobre CS y RA en el ámbito militar
- Estudiar y analizar las características funcionales de la RA para que sean aplicadas sobre dispositivos móviles utilizados por personal militar en el proceso de obtención de información circundante en un terreno desconocido y hostil.
- Diseñar e implementar un framework de software específico que tome como funcionalidades principales lo aportado por personal militar cumpliendo con las necesidades operativas propuestas.
- Desarrollar y probar artefactos desarrollados teniendo en cuenta el framework propuesto.
- Evaluar el cumplimiento de la CS mediante el uso de distintas técnicas de evaluación donde se contemple la experiencia del usuario
- Generar el conocimiento necesario en el tema para posteriormente transferir al personal militar las experiencias adquiridas.

1.3 Hipótesis

La principal hipótesis del presente trabajo de tesis radica en afirmar que la CS puede mejorarse en ámbitos militares con un uso adecuado de la RA. Para ello se apuesta por desarrollar un framework basado en la consideración de una arquitectura de software

integradora de tecnologías RA y requisitos tanto funcionales como no funcionales en el contexto considerado.

1.4 Metodología de investigación

Este trabajo de tesis se centró en la utilización de una metodología a fin de ordenar el proceso de investigación, diseño, pruebas y conclusiones finales. Para lograr este objetivo se optó por utilizar el framework de investigación DSR. En [7] se detalla que DSR se utiliza como un marco para guiar el proceso de investigación utilizando sistemas de información. El conjunto de directrices propuestas ha sido utilizado para planificar y ejecutar las actividades de investigación y desarrollo realizadas para este proyecto de doctorado.

Adicionalmente, se menciona que dicho proceso, en términos generales, puede seguir, más o menos, los siguientes lineamientos:

- Se identifica y se formula el problema.
- Se diseña un artefacto con el propósito de resolver partes o la totalidad del problema.
- El artefacto creado se evalúa utilizando métodos adecuados.
- El conocimiento adquirido al diseñar y evaluar el artefacto debe ser difundido.

Las partes y el detalle del workflow mencionado anteriormente se puede encontrar en la sección *10.1 Conclusiones*.

1.5 Organización del informe

El desarrollo del presente trabajo abarcó varios años de investigación y desarrollo y fue dividido en etapas con el fin de lograr el objetivo final propuesto. A continuación, se describen cada una de estas etapas, compuestas por partes, y los capítulos asociados a las mismas.

La *Parte I* de la tesis está conformada por el Estado de Situación del trabajo de investigación, consta de 5 capítulos cada uno analizando el estado actual de las áreas de conocimiento principal que conforman el tema de investigación.

En el Capítulo 1 se detallan la motivación del tema seleccionado en la presente tesis, los objetivos, la metodología de investigación utilizada, los aportes realizados, la hipótesis del trabajo de tesis, las publicaciones presentadas que surgen del proceso de investigación y desarrollo del presente trabajo de tesis y, finalmente, la organización del presente informe.

En el Capítulo 2 se realiza un estudio de la CS y de los antecedentes de proyectos militares que utilizaban componentes digitales. Adicionalmente, se estudian proyectos militares actuales que utilizan componentes tecnológicos digitales y son la antesala de los proyectos militares que utilizan técnicas de RA.

En el capítulo 3 se investiga en profundidad los fundamentos de la RA. Dentro de los fundamentos se detallan algunas definiciones de RA, los métodos de entrada de un sistema de RA, las partes que componen a un sistema de RA, los sistemas de coordenadas utilizados en un sistema de RA, el proceso de tracking, el proceso de reconocimiento de objetos y gestual, el modelo de referencia y la arquitectura de software utilizadas para el desarrollo de aplicaciones de RA.

En el Capítulo 4 se da cuenta del uso de RA para mejorar la CS. Se estudian proyectos militares que incorporan técnicas de RA para soportar la CS. Se realizan tablas comparativas de los proyectos digitales desarrollados por el Ejército de EEUU, tomando como referencia sus características principales a fin de evaluar si cumplen con lo solicitado por el personal militar del EE.AA.

En el Capítulo 5 se hace una introducción al uso de framework y su importancia. Se realiza una tabla comparativa de framework de RA comerciales (no militares) más comunes. Se detalla la arquitectura por capas y componentes que se tomará como base para el diseño del framework propuesto. Se especifica el ambiente de desarrollo sugerido y las técnicas de optimización adecuadas para el procesamiento de imágenes.

En la *Parte II* del trabajo de tesis se detalla el Framework Propuesto. Esta segunda parte está conformada por 5 capítulos.

En el Capítulo 5 se detalla el esquema conceptual de operaciones militares y la identificación y análisis de requisitos funcionales.

En el Capítulo 6 se detalla el esquema conceptual de operaciones militares aportadas por el EE.AA. Se identifican y se analizan en detalle los requisitos funcionales.

En el Capítulo 7 se especifica la metodología de desarrollo utilizada para el diseño de las aplicaciones de prueba. Se diseña la arquitectura del middleware del framework propuesto. Se detallan las capas y las librerías que forman parte del middleware. Se explica la arquitectura del sistema y los componentes de hardware en donde se ejecutarán los despliegues desarrollados. Se detalla cómo funciona el proceso distribuido implementado.

En el Capítulo 8 se da cuenta del despliegue de la solución que posteriormente se probará en las evaluaciones. Se detallan los dispositivos utilizados y los despliegues considerados.

En el Capítulo 9 se presenta la evaluación de los despliegues desarrollados en el capítulo anterior. Se evalúa la CS mediante la utilización de métodos válidos y confiables tales como SAGAT, SART, User Testing y Thinking Aloud y la experiencia del usuario mediante el cuestionario UMUX.

En el Capítulo 10 se detallan las conclusiones arribadas luego de la evaluación y experimentación realizada en el capítulo anterior. Adicionalmente, se mencionan las principales líneas de trabajo futuro que pueden seguirse como un estado evolutivo del framework propuesto.

Al final del documento se encuentra la bibliografía referenciada a lo largo del trabajo de tesis.

En la *Parte III* se encuentran los Apéndices correspondientes a la tesis doctoral.

Capítulo 2

Conciencia Situacional en proyectos militares

2.1 Introducción

En el presente capítulo se detalla qué es la CS. Se analizan los antecedentes de proyectos militares que utilizan componentes digitales para mejorar la CS en operaciones hostiles. Se evalúan diferentes tecnologías en el ámbito militar para la visualización y obtención de información contextual.

2.2 ¿Qué es la CS?

Brown, David Wm [2] menciona que la CS (en inglés, situation awareness o también situational awareness) se refiere a la percepción, la comprensión, y la previsión de los elementos dentro de un entorno operacional requerido para actuar con eficacia dentro de ese ambiente.

Tremblay, Sébastien et al [8] definen que la CS es un requisito previo para la oportuna y correcta toma de decisiones en el rápido y altamente estresante contexto de los entornos operativos de infantería. Se espera que la introducción de las tecnologías de soporte electrónico en el campo de batalla mejore la CS, proporcionando la información correcta, en el momento adecuado y en el formato correcto.

En la siguiente figura se grafica el modelo de la CS en la toma de decisión dinámica.

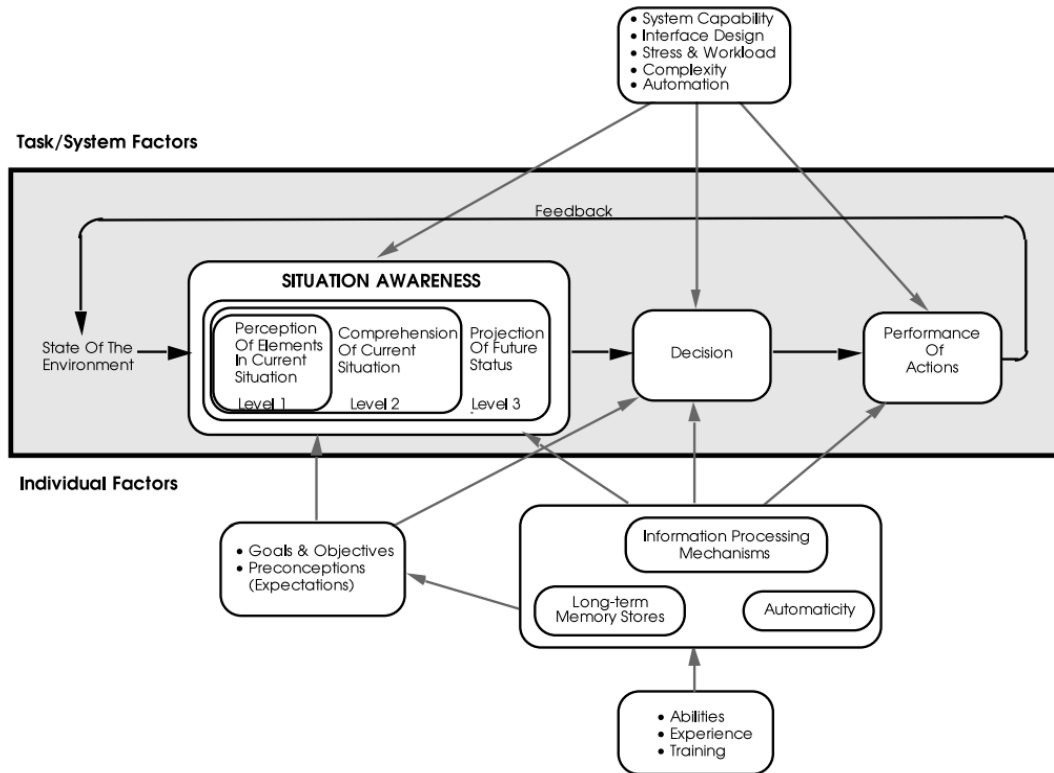


Figura 2.1 – Modelo de la CS en la toma de decisión dinámica (fuente: Masoud, G. et al [9])

Endsley, M. R. et al [10] determinan que uno de los factores más importantes que subyacen en el desarrollo de una adecuada CS es la presencia de los modelos mentales y esquemas de situaciones prototípicas. Proporcionan una construcción mental fundamental para dirigir la forma de asignar la atención y destacar los temas críticos.

Los principales atributos de la CS son:

- Visión global en todo momento con la posibilidad de obtener detalles en cualquier momento
- El uso de interfaz estándar para el acceso y el intercambio de datos
- Los datos organizados para alcanzar los objetivos
- Fácil transición entre los objetivos

Endsley, M. R [11] [12] menciona que la CS es la percepción de los elementos en el medio ambiente dentro de un volumen de tiempo y espacio, la comprensión de su significado y la proyección de su situación en un futuro próximo. Es importante detallar los tres niveles de la CS propuestos. El modelo sugerido se estructura en tres niveles jerárquicos de la CS,

cada uno de los cuales es necesario, pero no suficiente, precursor del nivel siguiente y superior. El mismo se basa en el establecimiento de una cadena de procesamiento de la información que, dividida en tres niveles, comienza con la percepción de la misma, prosigue con la interpretación y finaliza con la predicción.

Los procesos están formados por tres niveles aplicables a la CS, estos son:

- *Nivel 1: **percepción de los elementos del entorno.*** Se refiere a la percepción de la información recibida a través de la instrumentación disponible, del comportamiento del entorno, de la información asociada a otros sujetos del entorno, a amenazas circundantes, en definitiva, al entendimiento del ambiente. En este estadio no se interpretan los datos, sólo se almacenan en forma “bruta”, sin elaborar. Algún dato puede servir para confirmar el estado de una variable, pero sin ser integrado todavía por el individuo.
- *Nivel 2: **comprensión de la situación actual.*** Se alcanza a través de la percepción de los elementos del entorno. En este estadio no es necesario que se produzca una comprensión relevante para el desarrollo de su tarea por parte del operador. Es el nivel en el que se produce el proceso de integración de los datos. La CS va más allá de la percepción, contemplando también el proceso mediante el cual los sujetos combinan, interpretan, almacenan y retienen información. De esta manera la CS incluye no solo la percepción de la información del entorno, sino también la integración de múltiples piezas de información y la determinación de su relevancia en función de los objetivos del sujeto.
- *Nivel 3: **predicción del estatus futuro.*** Este es el nivel más elevado de la CS, que se asocia con la habilidad para proyectar el estatus futuro de los elementos del entorno. La precisión en la predicción depende enormemente de la información obtenida en los niveles 1 y 2. La anticipación proporciona al sujeto el tiempo suficiente para resolver conflictos y planificar su actuación para alcanzar sus objetivos.

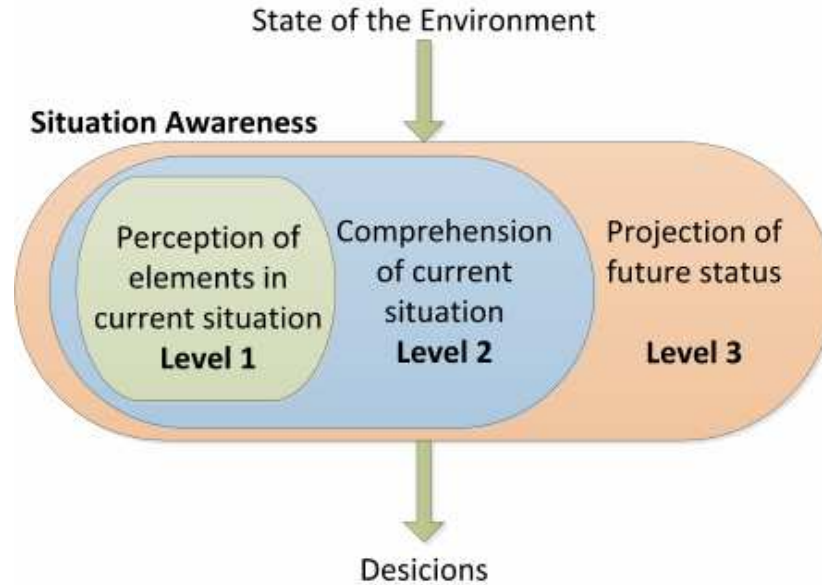


Figura 2.2 – Modelo de los tres niveles de la CS
(fuente: Endsley, M. R. [11] [12])

Por otro lado, los problemas más frecuentes asociados con la mala CS están dados por la ausencia en la recolección de la información necesaria, los problemas de percibir la información clave en la situación y la falta de comunicación. Las soluciones de CS permiten a los soldados hacer un uso efectivo de la información variada en un contexto de batalla.

De acuerdo a [13], la nueva tecnología ofrece métodos innovadores de obtener información contextual y representarla visualmente de una manera natural y no invasiva sin afectar el proceso cognitivo del combatiente.

2.3 Antecedentes de proyectos militares

Zieniewicz, Matthew J. et al [14] mencionan que en el año 1989, el Ejército de EE.UU. utilizó una pequeña computadora portátil para ayudar a los soldados en las tareas de campo de batalla. El concepto ha crecido desde entonces desde prototipos preliminares hasta la implementación del programa Land Warrior que incluye propuestas de futuros sistemas.

James Schoening, analista de investigación que trabajó en el CECOM (Communications Electronics Command) del Ejército de EE.UU., es quien comienza a utilizar computadoras portátiles. Trabajando con Matt Zieniewicz, Schoening transformó su idea en una arquitectura de sistema con tecnologías específicas, tales como la transmisión inalámbrica de

datos, captura de imágenes y GPS. En 1990, Schoening y Zieniewicz se asociaron con John Flatt, Sal Barone y Almon Gillette para demostrar el sistema Soldier's Computer. Más tarde, basándose en el proyecto "Soldier's Computer" (Figura 2.3), dio origen al proyecto SIPE (Soldier Integrated Protective Ensemble). El proyecto SIPE, dirigido por Carol Fitzgerald, fue el primero en que el Ejército de EE.UU trató a los diversos componentes de los equipos de combate como un sistema integrado. El equipo de diseño tuvo como objetivo desarrollar una computadora portátil con batería portable y software del campo de batalla. El equipo necesario debía incluir captura de imágenes, una radio integrada para la transmisión de datos entre los soldados y una unidad de visualización portátil, preferentemente un casco HMD.

El equipo decidió desde el principio evaluar los mejores componentes comerciales (en inglés COTS, Commercial Off-The-Shelf) en cada área (captura de vídeo, GPS, comunicaciones de datos, software de red, medios de almacenamiento, sistemas operativos, lenguajes de programación, procesadores) para llegar a la mejor arquitectura del sistema.

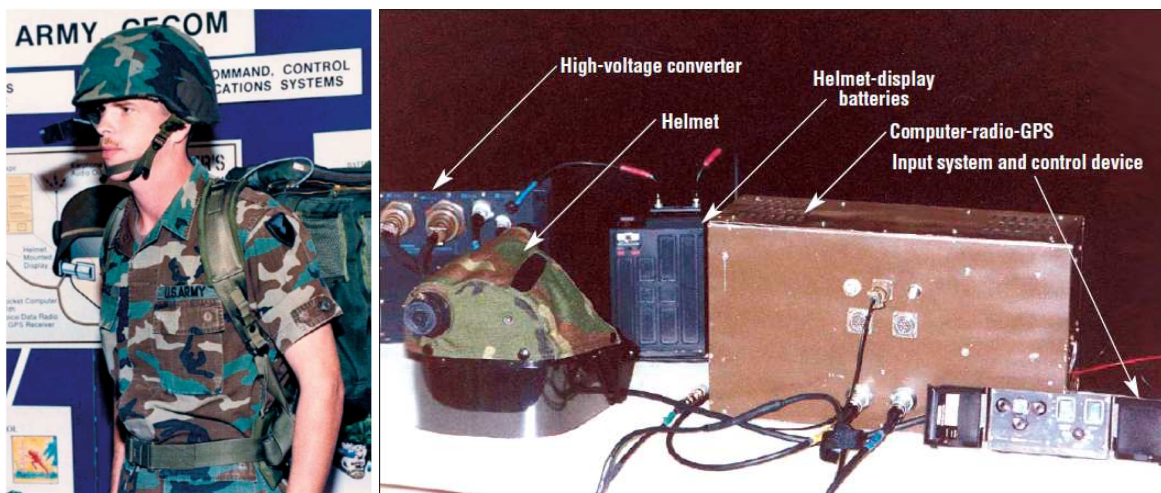


Figura 2.3 – Prototipo de Soldier's Computer junto con sus componentes electrónicos (fuente: Zieniewicz, M. et al [14])

- El sistema estaba dirigido a digitalizar las operaciones básicas para ayudar a los soldados:
- Leer mapas, navegar y mantener la conciencia situacional (para responder a las preguntas, por ejemplo, "¿Dónde estoy, ¿dónde están los miembros de mi equipo, y en qué dirección me estoy dirigiendo?")
 - Recibir, preparar y enviar informes escritos de campo (por lo que podría, por ejemplo, enviar una orden operacional, o preparar informes puntuales, etc)

- Capturar y transmitir imágenes fijas en color con fines de reconocimiento
- Acceder a material de referencia de las operaciones del campo de batalla (por ejemplo, las siluetas de los vehículos enemigos, los procedimientos de primeros auxilios, tareas comunes del campo de batalla, y los procedimientos estándar)

Aunque el sistema mejoró la capacidad de combate del soldado, el sistema tenía que ser más compacto y las baterías deberían operar más tiempo antes de que esté listo para el campo de batalla. Más importante aún, tenía que ser más liviano. Otro inconveniente fue el retraso en la captura y el envío de una imagen de vídeo, debido a las velocidades limitadas de procesamiento de la tarjeta de captura de vídeo y la capacidad del canal de comunicaciones. Como respuesta a estos inconvenientes surge el proyecto “Twenty-First Century Land Warrior”. Dicho proyecto examinó productos electrónicos avanzados, se basó en las tecnologías de comunicaciones, comando y control, informática, inteligencia, vigilancia y reconocimiento (en inglés, Command, Control, Communications, Computers, Intelligence, Surveillance and Reconnaissance -C4ISR-). En el año 1999, el equipo de diseño del Land Warrior del gobierno de EE.UU y el contratista asignado, comenzaron a trabajar en el primer diseño resistente del Land Warrior, bautizado con la versión 0.6. Se realizaron diferentes misiones con éxito. El objetivo del programa consistía en mejorar la eficacia de la unidad de infantería a través de tecnología y la integración de sistemas para funcionar como un sistema cohesivo. La integración se centraba en combinar los componentes del subsistema de armas con acceso a información visual y acústica provista por los sensores, añadiendo la capacidad de identificar la ubicación y posición, capacidades de recibir y transmitir datos y voz para el intercambio de información, determinar su propia ubicación para navegar y aumentar la protección balística y ataques QBR (Químico, Biológico y Radioactivo). Dado el éxito de la misión, se comenzó a diseñar la versión 1.0 del “Land Warrior”, pasándose a llamar “Land Warrior - Initial Capability -LW-IC”. En 2003, se desarrolló una variante del sistema LW-IC que incorporaba características de identificación de aliados y enemigos. Fue diseñado para reducir el potencial de incidentes de fuego aliado. Esta versión, se denominaba “Land Warrior – CombatID System -LW-CIDS-).

La necesidad de reabastecerse continuamente (o recargar) baterías estaba demostrando ser un gran desafío logístico. Para paliar este problema se centró en aquellas unidades terrestres que utilizaban vehículos del tipo Stryker. Este enfoque permitiría recargar las

baterías de los componentes electrónicos que formaban parte del “Land Warrior”. Esta versión pasó a llamarse “Land Warrior – Stryker Interoperable –LW-SI-).

El sistema de Land Warrior se compone de 7 subsistemas integrados:

- Subsistema de Casco Integrado (en inglés, Integrated Helmet Assembly Subsystem – IHAS-)
- Subsistema de Computadora
- Subsistema de Radio
- Sistema de Software
- Sistema de Armamento Modular (en inglés, Modular Weapon System –MWS-)
- Subsistema de Equipamiento Individual y Ropa de Protección (en inglés, Protective Clothing and Individual Equipment Subsystem)
- Subsistema de Navegación

Es importante destacar las características del subsistema IHAS (Figura 2.4). El subsistema está formado por un HMD con un visor monocular e intensificador de imagen de tercera generación con pantalla plana integrada. El IHAS funciona como interfaz con los otros subsistemas siendo éste el sistema de información visual primaria, pudiendo visualizar gráficos, mapas digitales, informaciones de misión, localización de tropas, información que provee el sistema MWS, entre otros datos.



**Figura 2.4 – Concepto del funcionamiento del subsistema IHAS
(fuente: Raytheon)**

Con el Land Warrior mejoró en gran medida la eficacia de las unidades de infantería en combate. La mejora en la Conciencia Situacional fue considerada un multiplicador de fuerzas. Una de las principales características del sistema “Land Warrior” incluía proveer al soldado de infantería información de aliados y enemigos para mejorar su Conciencia Situacional.

En paralelo con la última generación del programa “Land Warrior”, se desarrolló el programa “Objective Force Warrior”. Este programa buscaba demostrar como las tecnologías integradas apoyaban al combate individual. Incluía sistemas de armas, ropa protectora ligera con sensores integrados, comunicaciones de red, conciencia situacional cooperativa, navegación mejorada y monitoreo de estado físico. La prioridad máxima del programa estaba centrada en mejorar la Conciencia Situacional del grupo de combate a través de una red de comunicaciones. Para dar soporte a esto, se diseñó un casco del tipo HMD integrado con sensores conectado en red con vehículos terrestres y aéreos no tripulados (Figura 3.4).

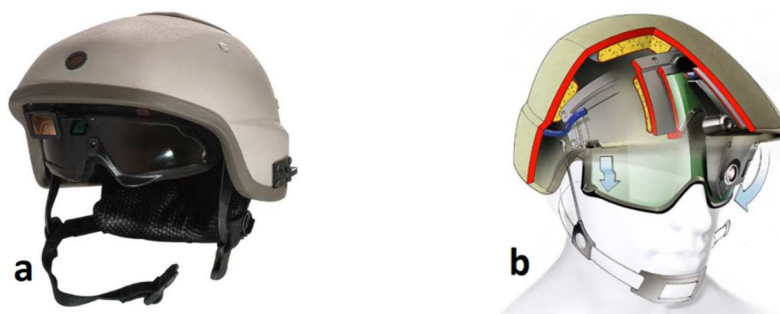


Figura 2.5 – (a) Primer prototipo funcional del HMD creado por Crye Associated – (b) Diseño conceptual del HMD con visión térmica, cámara de video, sensores QBR y un HUD para mostrar los datos e imágenes (fuente: Zieniewicz, M. et al [14])

2.4 Proyectos militares actuales

Los ejércitos de la mayoría de los países desarrollados están llevando acabo, desde hace tiempo, diversos programas de modernización destinados a mejorar las capacidades del soldado. Todos ellos coinciden en un denominador común, incrementar la eficacia y eficiencia del combatiente, proporcionando así mayor seguridad y capacidad de supervivencia en combate. Dichos programas aspiran a alcanzar un mismo objetivo: convertir al soldado de Infantería en un sensor de inteligencia y un órgano destinado a la adquisición de objetivos durante las operaciones, manteniendo igualmente la función esencial de ser una plataforma de armas, con capacidad de supervivencia en todas las situaciones. En la base de los diversos programas se encuentran las características de los conflictos actuales, asociados a las previsiones de cómo serán las confrontaciones bélicas del futuro.

Los programas dirigidos a la modernización de sus ejércitos de infantería conllevan un cambio revolucionario para la mayoría de los países avanzados. Los principales países de Europa que componen a la OTAN y EE.UU., encabezan los esfuerzos destinados a plasmar e industrializar esos avances tecnológicos.

Diferentes programas se están desarrollando en los ejércitos de diversos países, pero quizás los programas de Francia (FELIN), Alemania (IdZ - GLADIUS), Inglaterra (FIST) y España (COMFUT) sean los más consolidados por el tiempo que se ha invertido en el desarrollo de dichos programas y por centrarse objetivamente en mejorar la Conciencia Situacional.

FIST (Future Integrated Soldier Technology) es un programa de modernización de los soldados británicos que provee al infante con equipamiento de avanzada. El objetivo es crear un sistema de combate integrado para las tropas. El proyecto permitirá que el soldado del futuro esté equipado para maximizar su efectividad, mientras disminuye su carga física y psicológica, el efecto del estrés en combate y la posibilidad de error humano. En versión inicial del proyecto incluía un HDM que era considerado único pues utilizaba un sistema de imagen translúcido, basado en dos lentes transparentes en frente de los ojos brindando al soldado una mejora en la Conciencia Situacional. Se determinó que un equipo de comunicación en el grupo de combate y el uso de visión nocturna puede dar un aumento de rendimiento significativo. Por otro lado, la utilización de imagen térmica puede mejorar la Conciencia Situacional. La digitalización del FIST está dirigida a determinar qué información es importante para los soldados con el fin de mejorar la Conciencia Situacional, reduciendo la carga de trabajo del comandante y aumentando la coordinación y el ritmo operativo. El software de CS se llama Recognized Ground Picture (RGP).

Alemania, cuenta con el programa de combatiente futuro denominado IdZ (Infanterist der Zukunft). El IdZ básico, está destinado a los soldados de Infantería formados por grupos reducidos. El equipamiento, está constituido por un chaleco modular, que alcanza la protección balística y anti-esquirlas hasta el Nivel IV; equipo C4I, incluyendo NaviPad para control de situación y envío de mensajes, datos de blancos e imágenes, visor nocturno y radiotransmisor; y otros dispositivos especiales, como telémetro láser, cámara y dispositivos de designación por imagen de radiación térmica. Con él, no sólo se brinda protección a los soldados, sino que supone un notable aumento de las capacidades de mando y control,

navegación y empleo de armamento. El sistema IdZ pretende mejorar la Conciencia Situacional y la capacidad de respuesta con una mejor información sobre los enemigos y amenazas dando a todos los soldados del grupo de combate el mismo nivel de información. Como continuación del programa IdZ, se está desarrollando el proyecto Gladius para equipar al soldado de infantería. El equipo consta de diferentes componentes. Formado por un juego de chaleco antibalas, casco y gafas, visor infrarrojo óptico, computadora portátil, radio, equipo de navegación (GPS y brújula), así como armas y un conjunto de baterías. Adicionalmente, el equipo está diseñado para climas extremos y cuenta con sistema de protección contra armas biológicas y químicas.

El Ejército francés ha definido tres prioridades para el desarrollo futuro en el área de la guerra terrestre: la exploración del espacio de maniobra, supervivencia de combate, y la capacidad de moverse y pelear en las zonas urbanas. Con estos objetivos, el programa FELIN (Fantassin à Équipements et Liaisons Intégrés) desempeñará un papel clave. FELIN es un sistema altamente modular y totalmente integrado que traerá una mejora sustancial en todas las funciones operativas del soldado de infantería. Equipará a cada soldado con un sistema de combate completo, con una variedad de equipamientos electrónicos y optrónicos como visión diurna/nocturna, GPS, control de tiro automático, brújula, sistema de comunicación para transmisión de voz, datos e imágenes seguras. La miniaturización de los equipos y el suministro eléctrico asegurará que el soldado mantenga su agilidad y movilidad. Todos los equipos electrónicos que componen FELIN están relacionados entre ellos a través de una única computadora central, alimentados por baterías principales y de repuesto, ofreciendo una autonomía de 24 horas. Compuesto por gafas infrarrojas que permiten observar día y noche todos los blancos a una distancia de 300 metros, aumentando un 70 % el resultado de la visión diurna y 160 % la nocturna. El casco con equipo de comunicaciones basado en la osteofonía o resonancia de las placas óseas del cráneo, dejando las orejas libres. El soldado puede prestar mejor atención a los ruidos que le rodean manteniéndose comunicado.

El programa COMFUT (Combatiente Futuro) español nace con un claro objetivo: desarrollar un combatiente individual dentro de un pelotón que sea capaz de luchar en un campo de batalla digitalizado en cualquier conflicto futuro, y que no sólo actúe con una plataforma de armas sino también como un sensor de inteligencia y sistema de adquisición de objetivos. El diseño del sistema está integrado por diferentes componentes de modo que

trabajen en conjunto entre sí. Las partes del sistema incluyen una computadora portátil, un traje de protección, un casco integrado, armamento y sus sistemas de sustentación. La computadora es el cerebro del sistema y su software debe controlar todas las necesidades de los combatientes como información táctica y logística, vigilancia médica del soldado, sistema de puntería y visión nocturna conectado al arma del combatiente, reconocimiento de voz, gestión de información, GPS, radio y visores. El sistema de comando y control se basa en la idea de que el comandante del pelotón y los miembros del mismo tienen en todo momento una comprensión completa de dónde se encuentra cada combatiente, conociendo su estado operativo y proporcionando una información completa de la situación táctica en que se encuentra. En otras palabras, Conciencia Situacional.

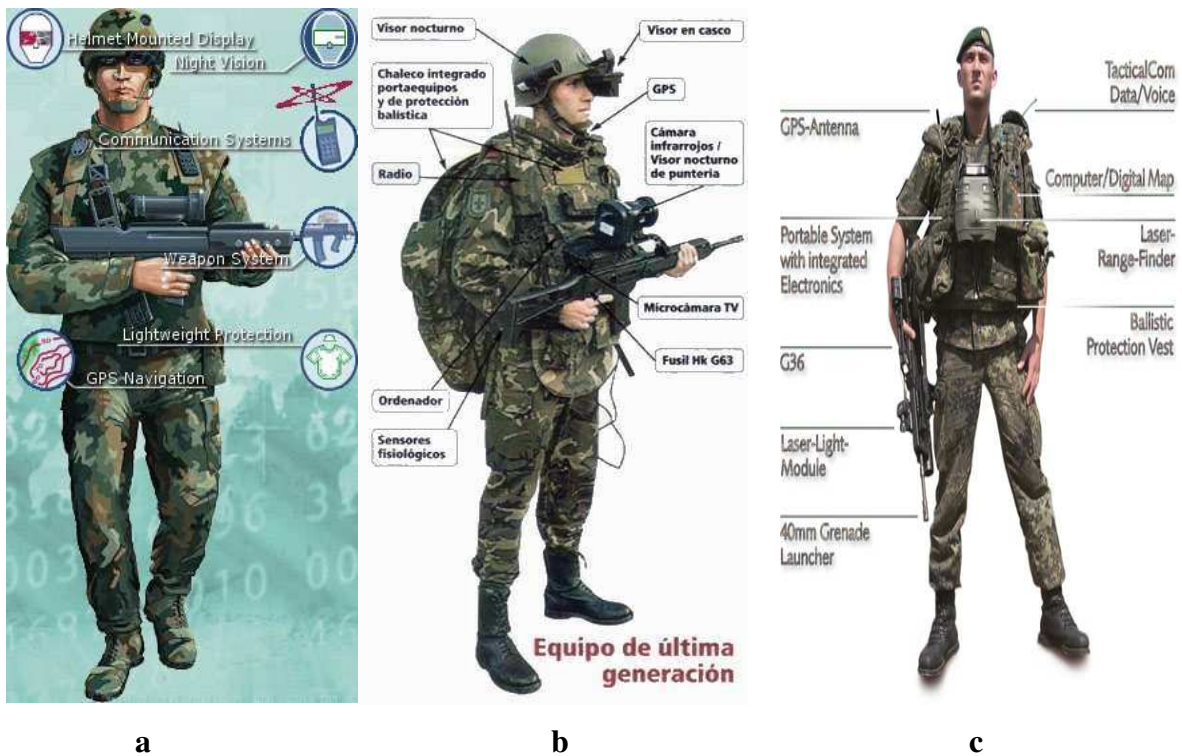


Figura 2.6 – Programas soldados del futuro: (a) FIST – (b) COMFUT – (c) IdZ (fuente: Zieniewicz, M. et al [14])

2.5 Resumen del Capítulo

Este capítulo presentó una introducción a la CS, mencionando a diferentes autores, destacándose el modelo de 3 niveles de Endsley: *Percepción* de los elementos en la situación actual, *Comprensión* de la situación actual y *Proyección* del estado futuro. Este modelo se

tomó como base para evaluar el despliegue integrador que se desarrolló a fin de analizar el cumplimiento y mejoramiento de la CS de los operadores en un ambiente simulado en una situación de estrés. Adicionalmente se analizan antecedentes de proyectos militares en la implementación de tecnología para mejorar la CS. Dichos programas forman parte de proyectos antecesores a la integración de tecnología utilizando, entre otras, la RA para mejorar la CS. Los proyectos donde se utiliza la RA como tecnología principal serán detallados y analizados en la sección 4.3 - *Realidad Aumentada en Proyectos Militares* .

Capítulo 3

Fundamentos de la Realidad Aumentada

3.1 Introducción

En este Capítulo se hace una introducción a los aspectos técnicos relacionados con la RA. Se menciona qué es la RA, las partes de una aplicación de RA como también los métodos de entrada de datos, visualización de información contextual y las tecnologías actuales para lograr este objetivo. Adicionalmente, se detalla qué es el tracking, el reconocimiento gestual y se hace mención de la arquitectura de software para aplicaciones de RA.

3.2 ¿Qué es la RA?

De acuerdo a [15], la Realidad Aumentada (RA) proporciona al usuario información superpuesta que se puede visualizar en el mundo real, es decir, complementa al mundo real con información virtual. La RA mejora la percepción del mundo natural mediante el agregado de información a los sentidos, ya sean visuales, sonidos, olores o sensaciones táctiles. La RA se refiere a la mezcla de las señales virtuales a partir del entorno tridimensional real en la percepción del usuario. Denota la fusión 3D de imágenes sintéticas en la visión natural del usuario del mundo circundante utilizando gafas, dispositivos móviles o un HMD (en inglés, Head-Mounted Display). Antes de continuar detallando en qué consiste la RA, es importante mencionar el uso de dispositivos móviles que implementan técnicas de RA. De acuerdo a [16] existen diferentes tipos de implementación de RA en dispositivos móviles. En este caso MAR (en inglés, Mobile Augmented Reality) puede ser entendido como cualquier

dispositivo móvil que utiliza RA. Estos dispositivos móviles pueden ser HDM, HMAR (en inglés, Handheld Mobile Augmented Reality), CAR (en inglés, Contact-lens augmented reality) o cualquier tipo de tecnología móvil que utilice RA para visualizar el mundo aumentado. Cabe destacar que HMAR es un subconjunto de MAR específicamente referido al uso de dispositivos handheld que pueden visualizar RA tal cual se muestra en la siguiente figura:

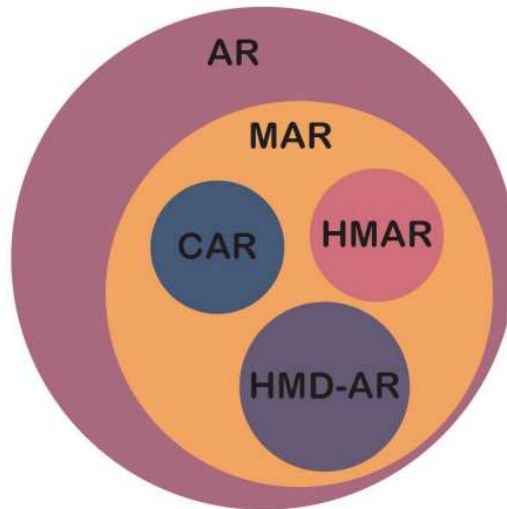


Figura 3.1 – La RA abarca un amplio espectro de sistemas. Dentro del dominio MAR Se encuentran sistemas tales como CAR, HMAR y HMD basadas en RA (fuente: Gjosaeter, T. [16])

La información espacial puede estar alineada directamente con el medio ambiente. A través de la capacidad de presentar la información superpuesta, integrada en el entorno del usuario, la RA tiene el potencial de proporcionar beneficios significativos en muchas áreas de aplicación. Muchos de estos beneficios surgen del hecho de que las señales virtuales presentadas por un sistema de estas características pueden ir más allá de lo que es físicamente visible.

Un sistema de RA es un tipo de sistema de visión sintética que mezcla gráficos generados por una computadora (anotaciones) con el mundo real. Las anotaciones proporcionan información destinada a ayudar a la toma de decisiones. El sistema de RA debe decidir qué anotaciones mostrar y cómo mostrar para asegurar que la información sea intuitiva y sin ambigüedades.

Un sistema de RA móvil consiste en un dispositivo móvil, un sistema de seguimiento (tracking) y un HMD, handheld o smartphone. El sistema hace tracking de la posición y orientación del usuario y superpone, dentro del campo de visión del usuario, gráficos y anotaciones que están alineados con los objetos en el medio ambiente. Este enfoque tiene muchas ventajas. La información puede ser presentada de una manera intuitiva y directamente integrada con el medio ambiente [17].

Cuando se implementa correctamente, la RA proporciona al usuario una experiencia de inmersión y permite la interacción entre el mundo virtual y el mundo real. Azuma [18] pone restricciones en la definición de la RA con la adición de tres criterios: combina lo real y lo virtual, interactiva en tiempo real, y están registrados y alineados en 3D.

En [19] se define a la RA como la tecnología de aumentación o ampliación de la percepción humana. Esta tecnología combina datos virtuales con el ambiente real observado por el usuario. Se produce la registración de información virtual con la escena real desde el punto de vista del usuario.

En [20] se introduce un término más genérico, realidad mixta, y lo definen como un *continuum* en donde uno de los extremos persiste el *ambiente real* y en el extremo opuesto persiste el *ambiente virtual*. La Realidad Aumentada y la Virtualidad Aumentada constituyen la realidad mixta o intermedia entre estos dos extremos.



Figura 3.2 – Continuum de la virtualidad
(fuente: Carmigniani, J. et al [5])

Azuma define tres requerimientos para las aplicaciones de RA:

- 1) Se deben combinar los aspectos real y virtual
- 2) La aplicación debe ser en tiempo real e interactiva, y
- 3) Los objetos deben estar registrados en el mundo tridimensional (3D)

Aunque el término RA deriva de la industria aeronáutica, otras industrias y aplicaciones rápidamente adoptaron el término. Los simuladores de vuelo se citan como ejemplos de RA.

No obstante, se considera a la RA a aquellas aplicaciones que combinan en una pantalla los escenarios reales y virtuales. Un buen ejemplo de representación visual de RA son los HMD. La información es proyectada de tal manera que un operador puede observar el mundo real directamente.

En el año 1992 Tom Caudell crea el término RA para describir una pantalla que usarían los técnicos electricistas de Boeing que mezclaba gráficos virtuales con la realidad física. La RA agrega información sintética a la realidad. La diferencia principal entre RV (Realidad Virtual) y RA es que por una parte RV implica inmersión del participante en un mundo totalmente virtual y por otra parte la RA implica mantenerse en el mundo real con agregados virtuales.

A diferencia de las aplicaciones de RV las aplicaciones de RA generalmente necesitan la movilidad del usuario, incluso hacia ambientes externos- en inglés se denominan aplicaciones outdoor. En dichas aplicaciones de realidad aumentada puede ser necesaria conocer la posición global del participante utilizando dispositivos como GPS y brújulas digitales.

Dependiendo del dispositivo de visualización usado las aplicaciones de realidad aumentada pueden basarse en:

- Gafas de *video see-through*
- Gafas de *optical see-through*
- Proyector
- Monitor
- Dispositivos móviles o HandHeld

En [21] se detalla que el ambiente real enriquecido o guiado con objetos virtuales puede resultar de mayor utilidad a los usuarios. Esta tecnología tiene diversas aplicaciones para la industria, la educación, el arte, etc. La principal diferencia con la realidad virtual es que no sustituye la realidad física, sino que superpone datos informáticos al mundo real.

El resultado final obtenido, permite al usuario el poder reconocer un objeto que este siendo apuntado en ese momento con la cámara del dispositivo y mostrar sobre éste, mediante realidad aumentada, información sobre el objeto. Específicamente la información extraída del escenario consiste de puntos característicos de la escena real detectados utilizando alguna técnica de tracking.

En términos de tracking, existen dos categorías de registros:

- Basados en sensores (mecánicos, magnéticos, ultrasónicos, etc.)
- Basados en visión por computador.

En los basados en sensores se deben calibrar los mismos y son pesados y caros y los niveles de precisión son pobres.

En cambio, los basados en visión por computador pueden evitar la calibración de sensores externos y ofrecen un muy nivel de precisión sin ser equipos pesados y costosos.

De acuerdo a la calibración de la cámara se puede mencionar:

- 1) No requiere calibración de los parámetros de la cámara (ya sean parámetros intrínsecos o extrínsecos), el cual involucra el uso de un objeto conocido de calibración 3D. Un método para resolver este problema es calcular los parámetros de la cámara y estructuras 3D en un espacio Euclidiano directamente de la información contenida en las secuencias de imágenes de entrada desde la cámara. Otra alternativa es ejecutar la tarea de RA en un espacio proyectivo o afín en vez de en un espacio Euclidiano.
- 2) Asume que los parámetros intrínsecos de la cámara están pre calibrados.

Bajo esta aproximación pragmática actual, para componer un servicio de RA son necesarios 4 ingredientes básicos [4]:

- Por un lado, un elemento que capture las imágenes de la realidad que están viendo los usuarios.
- Por otro, un elemento sobre el que proyectar la mezcla de las imágenes reales con las imágenes sintetizadas.
- En tercer lugar, es preciso tener un elemento de procesamiento, o varios de ellos que trabajen conjuntamente.
- Finalmente se necesita un elemento al que podríamos denominar «activador de realidad aumentada». En un mundo ideal el activador sería la imagen que están visualizando los usuarios, ya que a partir de ella el sistema debería reaccionar. Pero, dada la complejidad técnica que este proceso requiere, en la actualidad se utilizan otros elementos que los sustituyen. Se trata entonces de elementos de localización como los GPS que en la actualidad van integrados en gran parte de los Smartphone,

así como las brújulas y acelerómetros que permiten identificar la posición y orientación de dichos dispositivos, así como las etiquetas o marcadores del tipo RFID o códigos bidimensionales, o en general cualquier otro elemento que sea capaz de suministrar una información equivalente a la que proporcionaría lo que ve el usuario, como por ejemplo sensores.

De acuerdo a [22] hasta hace muy poco tiempo las aplicaciones de RA se ejecutaban en poderosas computadoras de escritorio. La introducción de RA en los smartphones permitió nuevas experiencias de RA a los usuarios. La RA se ha transformado en una comodidad ubicua para el ocio y para el aprendizaje. Con esta capacidad de ubicuidad, los móviles de RA permiten idear y diseñar escenarios de aprendizaje innovadores en entornos del mundo real.

Por otro lado, se puede definir a la RA como un sistema que mejora los sentidos primarios de las personas, es decir, visión, audición y táctil. Estos sentidos son mejorados o aumentados con información digital visible para el usuario. En este contexto, la información virtual incluye información como geo-localización, sobreimpresión visual/audio o información visual en 3D, entre otras.

Cabe destacar que al igual que los sistemas sensibles al contexto, las aplicaciones de RA permiten filtrar información y presentar información superpuesta relativa al contexto actual del usuario. La información del contexto puede ser filtrada de acuerdo a la ubicación, camino a recorrer, objeto en foco, período de tiempo o cualquier otra meta información tales como intereses o el perfil personal de la persona.

Por lo tanto, se define a la RA móvil como una especialización de los sistemas sensibles al contexto ya que existe una relación muy cercana entre los sentidos del humano y la percepción con la información digital presentada en RA.

De acuerdo a [23], los componentes que forman parte de un sistema de RA móvil deben contener:

- Sistemas de pantalla flexible, incluyendo un sistema HMD, cámaras en los Smartphone, y proyectores de mano.
- Sistemas de sensores en los dispositivos móviles tales como GPS, brújulas electrónicas, micrófono inclusive sistemas de tracking.
- Conexión a las redes de datos mediante tecnología inalámbrica.

- Poder computacional alto para el procesamiento de imágenes y visualización de objetos en 3D y sobreimpresión de información.
- Tecnología de tracking con 6 grados de libertad, tracking multi marcador y sistemas de tracking híbrido.
- Vinculación de la información de RA basada en localización
- Integración de navegadores de RA con los medios sociales.

La RA ha sido ampliamente demostrada y aplicada en dispositivos móviles, ejecutándose diversas aplicaciones tales como juegos, navegación, referencias, etc. Otro aspecto que está influyendo en las aplicaciones de RA es la tendencia a dirigirse hacia entornos donde se requiera la movilidad del usuario. Estas nuevas aplicaciones basadas en la computación móvil requieren acceder a servicios independientemente del lugar o del tiempo.

3.3 Partes de una aplicación de RA

Abásolo, M.J et al. [24], definen que una aplicación de RA tiene las siguientes partes:

- Captura de la escena real: El video capturado puede utilizarse para tracking basado en visión, es decir basado en el análisis de la imagen mediante algoritmos de visión
- Tracking del usuario: puede realizarse mediante dispositivos específicos o puede realizarse tracking basado en visión para lo cual es necesaria la captura de la escena real.
- Generación de la escena virtual: se tiene un mundo virtual, con la información de la posición y orientación del participante se genera una vista acorde del mismo
- Rendering: se combinan las imágenes del mundo real con los objetos virtuales. Los objetos vituales se renderizan y se proyectan en el dispositivo móvil.

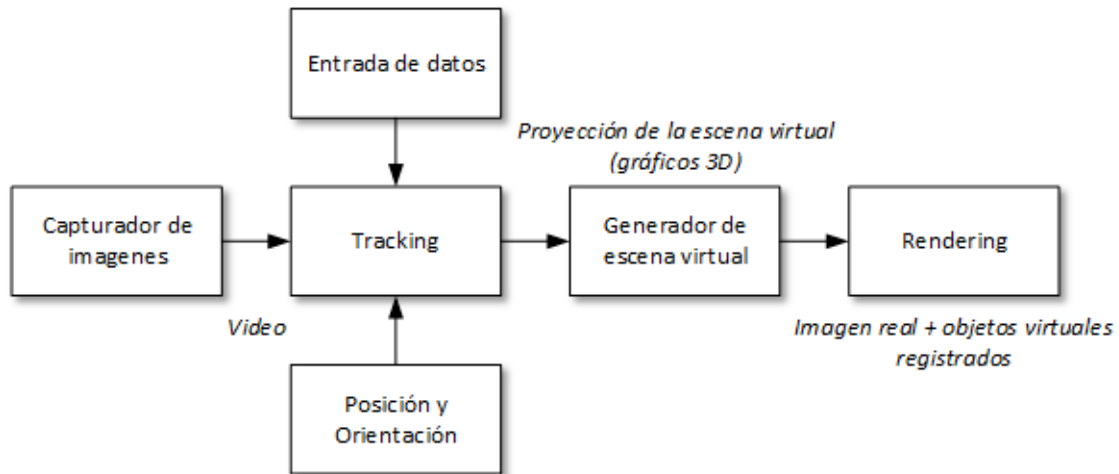


Figura 3.3 – Diagrama de las partes de una aplicación de realidad aumentada (fuente: propia)

Adicionalmente para el proyecto de tesis se agrega un componente más:

- Entrada de datos: se captura el video y se toma como entrada de datos el reconocimiento gestual o el reconocimiento por voz para ejecutar un comando determinado

Por otro lado en [18] especifica las partes o componentes principales de un Sistema de RA:

1. *Generador de Escena*: En los sistemas de RA las imágenes virtuales complementan al mundo real a diferencia de los entornos de RV que necesitan mayores requerimientos para las imágenes ya que reemplazan completamente al mundo real. En el caso de los sistemas de RA las anotaciones, textos, dibujos en 3D podría ser suficiente para representar cierta información.
2. *Dispositivo de visualización*: En los sistemas de RA el dispositivo de visualización puede ser menos rigurosos que los utilizados en RV ya que no reemplazan el mundo real, sino que lo complementa. Por ejemplo, pantallas monocromáticas puede ser útiles para algunos sistemas de RA, mientras que los sistemas de RV deben utilizar pantallas a color. Otro ejemplo son los HMD de visión óptica (*HMD optical see-through*), incluso pueden ser satisfactorios debido que el usuario puede ver el mundo real con su visión periférica. También se puede mencionar los dispositivos móviles tales HMAR, CAR, smartphone, entre otros.
3. *Tracking (seguimiento) y detección*: En este caso el tracking y la detección juega un rol importante en los sistemas de RA. En esta área los requerimientos para los sistemas de

RA son muchos más estrictos que los sistemas de RV. La principal razón es el problema de registración de la información.

Modelo de referencia

Un Modelo de Referencia en ingeniería de sistemas y en ingeniería de software es un modelo de algo que contiene un objetivo o idea básica de algo, y que se puede establecer como una referencia para múltiples propósitos.

También se puede definir como un marco de referencia abstracto para entender el significado de las relaciones entre entidades de algún ambiente. Permite el desarrollo de referencias específicas o de arquitecturas por medio del uso de estándares o especificaciones que soportan el ambiente en cuestión. Un Modelo de Referencia consiste de un conjunto mínimo de conceptos, axiomas y relaciones propios de un dominio particular de problema, y es independiente de estándares específicos, tecnologías, implementaciones, o de cualquier otro detalle concreto.

Un estudio de varios sistemas de RA reveló que, a pesar de ser bastante diferentes en los detalles, la mayoría de los sistemas de RA existentes comparten una estructura arquitectónica en común [25]. En base a esto, se diseñó una arquitectura de referencia [25] como la que se muestra en la siguiente figura:

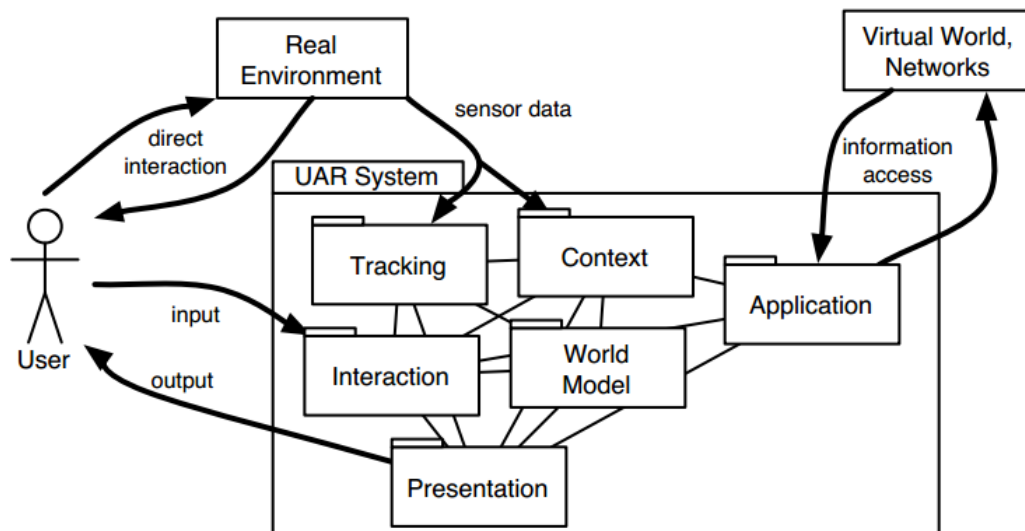


Figura 3.4 – Modelo de referencia. Componentes de software requeridos en un sistema de RA ubicua (fuente: MacWilliams, A. [25])

El Modelo de Referencia se organiza en seis subsistemas lógicos:

- *Subsistema de seguimiento (tracking)*: Responsable de responder a los cambios de posición y orientación del usuario.
- *Aplicación*: Responsable del control de flujo, la lógica de la aplicación y la coordinación.
- *Modelo del mundo*: provee acceso a la representación digital del mundo.
- *Presentación*: Sistema responsable de la presentación de los resultados. Incluye renderizado 2D/3D y salida de audio o táctil.
- *Contexto*: provee a la aplicación con contexto tanto del estado como de la situación del usuario.
- *Interacción*: representa la interacción del usuario con el sistema de RA.

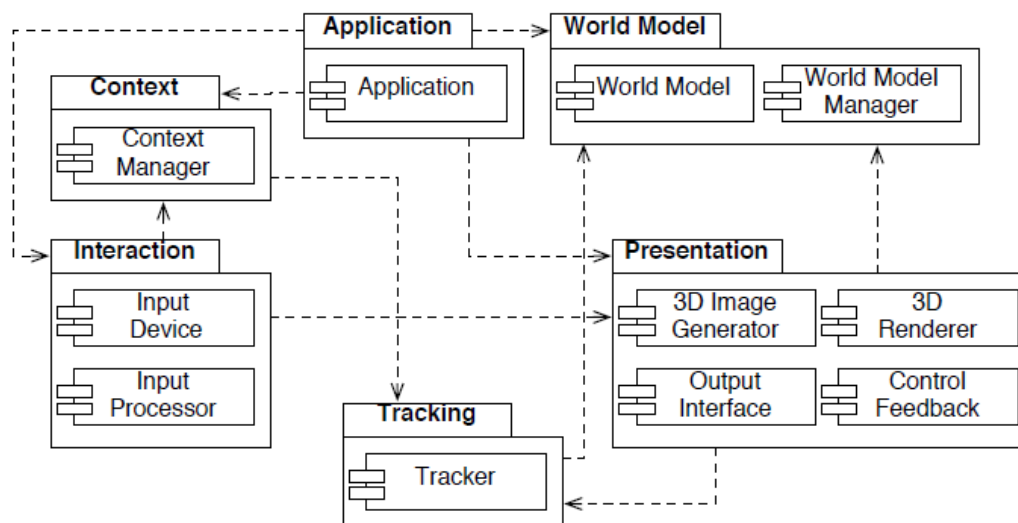


Figura 3.5 – Arquitectura de referencia genérica de sistemas de RA
(fuente: MacWilliams, A. [25])

Cada subsistema está formado por varios componentes. Los subsistemas se muestran como un paquete en UML lo cuales están compuestos por componentes. Por otro lado, cada componente a su vez puede ser obtenido por medio de otros componentes. Hay relaciones de dependencia entre los subsistemas. Una dependencia muestra que un subsistema se basa en las interfaces de otro subsistema.

El modelo de referencia describe a nivel abstracto los componentes generales y la estructura de los sistemas de realidad aumentada. Sin embargo, dependiendo de los requisitos funcionales de un sistema particular, algunos de los componentes pueden dejarse afuera [26].

3.4 Métodos de entrada de datos en un sistema de RA

Resulta importante entender los métodos de entrada de datos en un sistema de RA. Esta importancia se denota ya que poder ingresar comandos para una determinada acción sin utilizar, por ejemplo, el teclado se transforma en una posibilidad a tener en cuenta sobre todo cuando es primordial tener las manos libres de periféricos para el ingreso de información. En los sistemas de RA se pueden mencionar los siguientes mecanismos de toma de datos:

De acuerdo a [24] se establece los siguientes mecanismos de entrada:

- Controles físicos
- Control del habla
- Control por gestos

En un sistema de RA podemos encontrar dentro de los controles físicos a los dispositivos móviles, sean estos smartphones, tablets, etc. Por otro lado, el control del habla es muy adecuado cuando se tienen las manos ocupadas. Mientras que el control por gestos es un método natural de comunicar la información en el que se basan nuevas metáforas de interacción.

Requiere de dispositivos especiales para el tracking o en su lugar de la aplicación de técnicas de visión por computador.

Kölsch et al [27] mencionan la existencia de las interfaces multimodales donde se combinan, por ejemplo, reconocimiento de voz y gestos. Por lo tanto, se puede llegar a la siguiente conclusión:

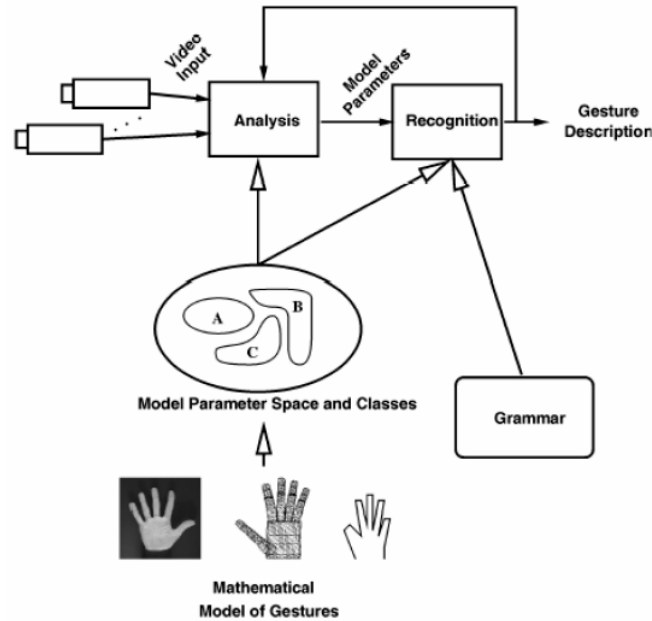
- Aquellas tareas que requieren parámetros adimensionales, es decir, no se debe indicar nada en el espacio, pueden indicarse adecuadamente mediante el habla
- Las tareas que requieren parámetros espaciales de una dimensión como un desplazamiento en un eje (por ejemplo, indicar un zoom)
- Pueden indicarse adecuadamente mediante un dispositivo como un trackball unidireccional.
- Las tareas que requieren parámetros espaciales de dos dimensiones y tres dimensiones (posicionamiento, escalado y orientación de objetos en el mundo virtual) requieren utilizar diferentes tipos de entrada combinando gestos, movimientos de cabeza y trackball unidireccional.

Hay muchos tipos de dispositivos de entrada los sistemas de RA. Algunos sistemas utilizan guantes para ingresar datos, otros utilizan algún sistema inalámbrico. En el caso de los smartphones pueden ser utilizados como dispositivos de orientación o posicionamiento como es el caso de Google Sky Map para el sistema operativo Android [5].

La interacción humana con la computadora ha sido por años una manera de comunicación centrada en las máquinas. El reconocimiento gestual es una forma natural de expresión humana, por lo humano-computadora (en inglés, Human-Computer Interaction – HCI-). Una solución es utilizar técnicas de visión para el reconocimiento gestual. La mayoría de las estrategias de reconocimiento gestual basadas en visión por computador se centran en reconocimiento de gestos estáticos conocido como posturas. Sin embargo, se ha argumentado que el movimiento de los gestos transmite tantos significados como la propia postura. Los movimientos incluyen, por ejemplo, movimiento global de la mano y análisis del movimiento del dedo [19].

La interpretación de los gestos puede dividirse en tres fases: modelado, análisis y reconocimiento.

- *Modelado*: El modelado de los gestos involucra la descripción esquemática de un sistema de gestos que da cuenta de sus propiedades conocidas o inferidas.
- *Análisis*: el análisis del gesto involucra el cálculo de los parámetros del modelo basado en las características de la imagen detectada por la cámara.
- *Reconocimiento*: esta fase involucra la clasificación de los gestos basado en los parámetros del modelo procesado.



**Figura 3.6 – Sistema de reconocimiento gestual
(fuente: McDonald, C. [19])**

Un acercamiento muy interesante es el que se menciona en [28] donde se detalla una interacción más precisa en el reconocimiento gestual en lo que se refiere al tracking para la interacción de la interfaz del usuario en RA dónde se evalúa la usabilidad y la precisión. Esta solución combina algoritmos para la extracción del contorno de la mano para aplicarlo en una aplicación de RA. De esta manera, se ocupa de los siguientes requisitos:

- Selección de los ítems del menú utilizando reconocimiento gestual
- Robustez y precisión en la detección de la mano y tracking en condiciones de variación de la iluminación con oclusión limitada
- Uso de la mano como un puntero en 3D

La facilidad de uso y la independencia de los gestos se abordan mediante la definición de cuatro posturas intuitivas de las manos que sirven para la navegación de los menús y la selección. Las tareas de navegación y selección son soportadas por detección automática de objetos de la mano y extracción de formas de la mano.

Una de las técnicas más populares y robustas utilizadas para la detección de la mano es la propuesta por Viola & Jones [29]. Este método utiliza Adaboost un método supervisado que ejecuta clasificación binaria. Durante el entrenamiento, se genera iterativamente un

clasificador denominado fuerte agregando un clasificador débil cuya función es adivinar al azar. Cada clasificador débil es entrenado en un conjunto de datos generado donde se aplica muestreo o ponderación en los casos del conjunto de datos original de la mano.

Active Appearance Model (AAM) es un modelo estadístico que maneja variaciones de formas y texturas de la apariencia foto realística. AAM utiliza un conocimiento previo de la apariencia en escalas de grises, estructura de la forma inclusive sus relaciones, con el objetivo de construir modelos generativos para un análisis global de una clase específica de objetos. Esta técnica se utiliza para recuperar formas en 2D de la postura de la palma y puño de la mano en un sistema de RA, siguiendo una serie de análisis locales, alrededor del área rectangular de la mano utilizando la técnica de Viola & Jones [29]. En RA cuando se utiliza el detector Viola & Jones de la palma de la mano para estimar la ubicación de la punta del dedo índice, y AAM para permitir un puntero basado en el dedo.

3.5 Visualización

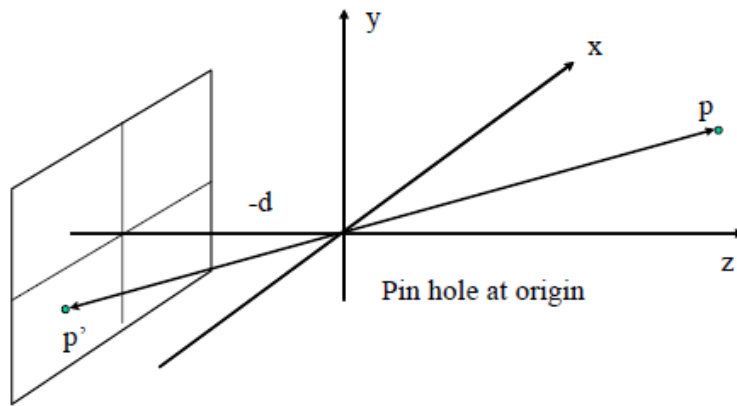
Sistemas de Coordenadas

El principal desafío para los sistemas de RA es cómo combinar el mundo real y el mundo virtual en un único entorno.

Como menciona [24] existen dos tipos de proyecciones más conocidas y utilizadas en informática gráfica: proyección ortogonal y proyección perspectiva.

La proyección perspectiva puede describirse por un modelo matemático sencillo denominado “modelo de Pinhole”. Este modelo de cámara es utilizado en computación gráfica y permite modelar la transformación proyectiva de una escena de tres dimensiones en un plano de dos dimensiones. En este modelo de cámara, el lente de la cámara (pin-hole) se encuentra en el origen y un punto p se proyecta sobre la película en el punto p'

La distancia entre la película fotográfica y el lente se conoce como longitud focal d [19]



Usando este modelo, se puede definir la relación entre las coordenadas 3D en la escena virtual, x e y , y el resultado de las coordenadas de la imagen en 2D, x' e y'

$$x' = d \frac{x}{z} \quad \text{and} \quad y' = d \frac{y}{z}$$

En su forma general, esta relación puede ser representada a través de la siguiente transformación homogénea:

$$p' = Mp,$$

donde p y p' son puntos homogéneos y M es la matriz de proyección 4x4:

$$\begin{bmatrix} x' \\ y' \\ z' \\ w \end{bmatrix} = \begin{bmatrix} d/z & 0 & 0 & 0 \\ 0 & d/z & 0 & 0 \\ 0 & 0 & d/z & 0 \\ 0 & 0 & 1/z & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

Para obtener esta matriz de proyección para una posición arbitraria de la cámara en el espacio, los parámetros intrínsecos y extrínsecos de la cámara se deben extraer independientemente.

Para mantener esta ilusión al usuario, de que los objetos virtuales en realidad son parte del mundo real, requiere una fusión coherente entre el mundo virtual con el mundo real.

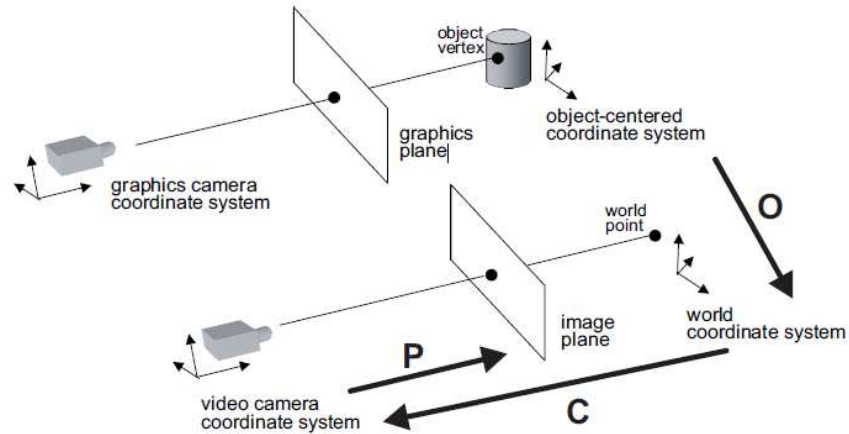


Figura 3.7 – Sistemas de Coordenadas de un sistema de RA
(fuente: McDonald, C. [19])

El principal requisito es determinar las relaciones entre ambos mundos. Estas relaciones son del sistema de coordenadas objeto a mundo O , de mundo a cámara C , y de cámara a una imagen plana P . De objeto a mundo se refiere a la posición y orientación de un objeto virtual con respecto al sistema de coordenadas del mundo que define la escena real. La de mundo a cámara define la posición de la videocámara que apunta hacia la escena real. Por último, de cámara a una imagen plana especifica la proyección que la cámara realiza para crear una imagen 2D a partir de la escena real en 3D.

La creación de una imagen tridimensional de un objeto virtual en consonancia con la vista actual del usuario en el mundo real y la colocación del objeto que está viendo en dicho mundo, requiere la definición de las relaciones geométricas entre los objetos virtuales y los físicos.

Como también se especifica en [19] el sistema de coordenadas del mundo es el punto de referencia inicial. Luego, el sistema de coordenadas del video cámara se debe determinar utilizando alguna estrategia de visión por computador. La transformación del sistema de coordenadas del mundo al sistema de coordenadas de la cámara está representada por C .

La transformación proyectiva definida por medio del modelo de cámara está representada por P . La transformación final necesaria para llevar a cabo la registración de manera adecuada es la transformación del sistema de coordenadas centrado en el objeto al sistema de coordenadas del mundo, O . Las coordenadas 3D del objeto virtual son asignadas a priori, por lo que esta transformación se construye en ese momento.

Con las dos cámaras alineadas, la fusión de los componentes reales y sintéticos de la escena será debidamente registrada.

De acuerdo a [24], en el proceso de formación de imágenes, reales o sintéticas, existe:

- una *escena* compuesta por diversos objetos (en diferentes posiciones y orientaciones)
- una *cámara*, con una cierta posición y orientación en la escena
- una *imagen* resultante de proyectar la escena de acuerdo a la cámara

Sistema de coordenadas local

Es un sistema de coordenadas 3D utilizado para referenciar los puntos de un objeto. Si se modela un árbol seguramente lo haremos situando un sistema de coordenadas centrado en el árbol o en la base del mismo, con un eje alineado con el tronco.

Sistema de coordenadas mundo

La posición y orientación de los objetos de una escena suele expresarse en relación a un sistema de coordenadas 3D situado en algún lugar del mundo. Todos los objetos que integran una escena se posicionan, orientan y escalan en relación al mismo sistema de coordenadas mundo.

Sistema de coordenadas cámara

El sistema de coordenadas cámara 3D tiene el origen en el centro óptico de la cámara. El eje de proyección de la cámara pasa por el centro óptico y es perpendicular al plano de formación de la imagen. El eje Z del sistema de coordenadas cámara está alineado con el eje de proyección, con Z+ hacia donde se ubica la escena. El eje Y tiene dirección vertical y el eje X dirección horizontal.

El sistema de coordenadas mundo 3D puede expresarse en relación al sistema de coordenadas cámara 3D mediante transformaciones geométricas. De forma equivalente, el sistema de coordenadas cámara puede expresarse en relación al sistema de coordenadas mundo.

Sistema de coordenadas de la imagen

Los puntos de la imagen se expresan en relación a un sistema de coordenadas 2D con los ejes alineados con los bordes de la imagen.

Es importante destacar lo que menciona [30] en lo que se refiere a los sistemas de coordenadas cuando se manipulan objetos 3D. Los autores evalúan el impacto de los tres sistemas de coordenadas: cámara, objeto y mundo, sobre la rotación y la traslación de un objeto 3D en un escenario de RA. Los ejes del Sistema de coordenadas de la cámara dependen de la pose actual del dispositivo móvil, también conocido como la vista del usuario. La vista del usuario se utiliza frecuentemente como sistema de referencia en los dispositivos móviles de RA. Cada objeto en el espacio tiene su pose 3D específica. La característica del sistema de coordenadas del objeto es la dependencia de los ejes entre la traslación y la rotación. El Sistema de coordenadas del mundo es el único sistema de coordenadas estático y representa el espacio en el mundo real.

Por otro lado en [5] se mencionan algunos métodos de visión por computador utilizados en RA. La registración de la imagen en RA utiliza diferentes métodos de visión por computador generalmente relacionado con el tracking de video. Estos métodos usualmente consisten de dos etapas: tracking y reconstrucción/reconocimiento. Primero, los marcadores fiduciaros, imágenes ópticas o puntos de interés son detectados en las imágenes de la cámara. El tracking puede utilizar métodos de procesamiento de imágenes tales como detección de características, detección de contornos u otros métodos a fin de interpretar las imágenes de la cámara.

En visión por computador la mayoría de las técnicas de tracking pueden separarse en dos clases: basados en características y basados en modelos. El primer método consiste de descubrir la conexión entre las características de la imagen en 2D y las coordenadas del mundo en 3D. El segundo método utiliza modelos de los objetos a trackear tales como modelos CAD o plantillas en 2D.

Una vez que se realiza la conexión entre la imagen en 2D y el frame del mundo en 3D, es posible encontrar la pose de la cámara proyectando las coordenadas 3D de la característica en las coordenadas de la imagen en 2D y minimizando las distancias a sus características correspondientes en 2D.

Las limitaciones de la estimación de la pose de la cámara son frecuentemente determinadas utilizando las características de los puntos.

La etapa de reconstrucción/reconocimiento utiliza los datos obtenidos de la primera etapa para reconstruir el sistema de coordenadas del mundo real.

Registro de información

La RA se ha transformado en una tecnología prometedora para soportar tareas en 3D. La registración es una de las claves más importantes de esta tecnología. El proceso de registro es la superposición de objetos virtuales en una escena real utilizando información extraída de la escena [31].

Como menciona [24] la información virtual tiene que estar vinculada espacialmente al mundo real de manera coherente, lo que se denomina registro de imágenes (“registration” en inglés). Por esto se necesita saber en todo momento la posición del usuario, tracking, con respecto al mundo real.

Cuando se habla de posición y orientación de un objeto o de una cámara debe existir un sistema de referencia en base al cual se expresan. En RA, al hablar de imágenes registradas significa que tanto las imágenes sintéticas como el mundo real estén en referencia al mismo sistema de coordenadas. Es importante destacar la registración geométrica de los objetos sintéticos en la escena real. Para ello [32] detalla que la registración geométrica del contenido virtual dentro del ambiente real del usuario es la base para una aplicación de RA. Una correcta registración geométrica le permite al observador aceptar el contenido virtual como una mejora del ambiente real más que una separada o superpuesta capa de información. La escena virtual se registra geoméricamente midiendo o estimando la pose actual del punto de vista del usuario o la cámara con respecto al sistema de coordenadas del contenido virtual.

3.6 Tecnologías para visualizar la RA

Azuma, R [18] menciona para construir un sistema de RA es importante determinar qué tecnología utilizar para combinar el mundo real con el digital. Para ello existen dos tipos de tecnologías: óptica y video. Un HMD *see-through* es un dispositivo que se utiliza para para combinar el mundo real con el digital. Este tipo de dispositivo permite ver el mundo real superimprimiendo objetos digitales a través del uso de la tecnología óptica o por video. La tecnología óptica trabaja combinando objetos digitales en frente del campo visual del usuario. Generalmente se utiliza un combinador óptico transparente del tipo Head-Up Displays (HUDs) como se utiliza en los cascos de los pilotos de aviones de combate. El combinador óptico, usualmente reduce la cantidad de luz que usuario visualiza en el mundo real para

lograr una alta absorción de la luz se utilizan proyectores del tipo monocromáticos o gafas oscuras donde directamente reducen la onda de luz proveniente del mundo real.

En la siguiente figura se muestra el diagrama conceptual del uso de la tecnología *optical see-through* utilizando un HMD:

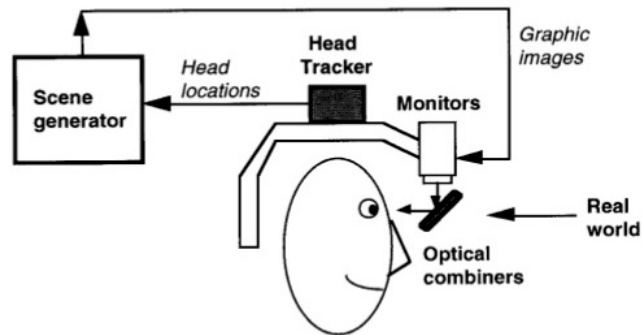


Figura 3.8 – Diagrama conceptual de la tecnología *optical see-through*
(fuente: Azuma, R [18])

Por otro lado, la tecnología *video see-through* trabaja combinando, en un HDM cerrado, una o más video cámaras (monocular o binocular, respectivamente) montadas en el casco. Los videos cámaras proveen al usuario la vista del mundo real. El video de estas cámaras se combina con las imágenes gráficas creadas por el generador de escenas, mezclando el mundo real con el digital. El resultado se envía a los monitores (o pantallas) en frente de los ojos del usuario. En la siguiente figura se muestra el diagrama conceptual de la tecnología *video see-through*:

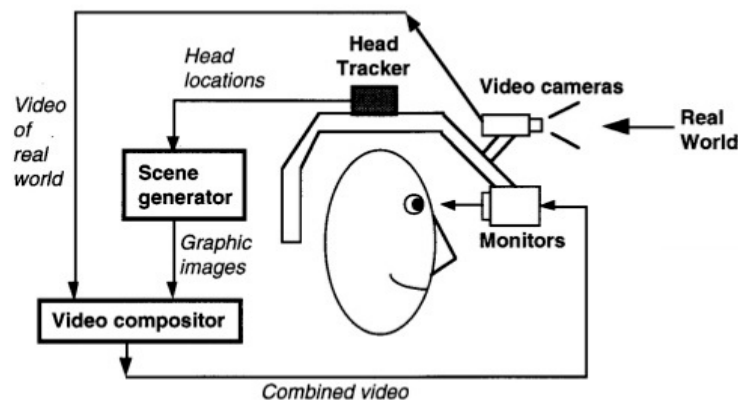


Figura 3.9 – Diagrama conceptual de la tecnología *video see-through*
(fuente: Azuma, R [18])

3.7 Tracking

En [24] define al tracking (seguimiento) como el proceso de seguimiento de un objeto en movimiento, es decir, la estimación de la posición y orientación (pose) del mismo en cada instante.

En una aplicación de RA se necesita el tracking del participante para conocer la matriz de transformaciones geométricas y así realizar el registro de imágenes sintéticas y reales.

En la mayoría de aplicaciones de RA se trata de realizar un tracking de la cámara que captura la escena. Aunque también puede tratarse del tracking de la cabeza del usuario o de algún objeto manipulado por este.

Los autores también mencionan que el tracking en una aplicación de realidad aumentada puede hacerse:

- mediante dispositivos físicos específicos
- tracking basado en visión
- tracking híbrido, que combina las salidas de dispositivos físicos con el análisis de la imagen

Por otro lado [5] menciona que aunque el tracking visual tiene la habilidad para reconocer y seguir varias cosas, mayormente utiliza tecnologías tales como GPS y acelerómetros.

También en [33] menciona métodos de visión por computador se utilizan para el tracking de RA. A diferencia de otras tecnologías, los métodos de visión pueden estimar la pose directamente de la misma imagen que el usuario observa.

En [5] detalla que los métodos de tracking en RA dependen mayormente del tipo de ambiente que el dispositivo móvil se encuentre inclusive el tipo de sistema de RA. El ambiente puede ser interior, exterior o una combinación de ambos. De la misma manera, el sistema puede ser móvil o estático (con una posición fija). Si el dispositivo de RA es móvil y se diseñó para ambientes exteriores. El tracking será mucho más duro y dependiendo de las técnicas implementadas ofrecerán algunas o desventajas.

El tracking por visión por computador es eficiente para calcular la pose de la cámara en relación a la escena objetivo. Si bien el poder de hardware de los dispositivos móviles aún

no es del todo poderoso se han explorado técnicas de tracking mediante características naturales.

En [34] menciona que el tracking de objetos mediante visión por computador es uno de los tópicos de investigación más activos en visión por computador por su importancia en teleconferencias, vigilancias, seguridad, interacción humano computadora, comprensión de video y edición de video, etc.

La tarea de encontrar correspondencias entre las imágenes de una misma escena o el objeto es esencial en las aplicaciones de visión por computador. Esto puede ser implementado utilizando tres pasos: detección, descripción y comparación (matching)

En el paso de detección se seleccionan los puntos de interés de distintos lugares de la imagen tales como esquinas y bordes. Estos puntos de interés deben ser distintivos y repetibles, es decir, deben poder detectarse bajo diferentes condiciones visuales.

En el paso de descripción, los puntos vecinos de cada punto de interés se representan por medio de un vector de características. Este proceso debe ser robusto a ruidos, detección de errores y deformaciones geométricas y fotométricas.

Por último, en la etapa de comparación, los vectores de características de las diferentes imágenes son comparadas. Esto se hace usualmente calculando la distancia Euclidiana entre los vectores de características.

En [35] recientes técnicas de comparación (matching) de imágenes tal como BRIEF reducen el tiempo omitiendo el cálculo del factor invariante de la escala, y en cambio, se considera variaciones multiescala de la imagen destino en el pre-procesamiento.

En [36] determina que el tracking de objetos es pertinente en las tareas de:

- Reconocimiento basado en movimiento, es decir, identificación humana basada en el paso, detección automática de objeto, etc.:
- Vigilancia automatizada, es decir, monitorear una escena para detectar actividades sospechosas;
- Indexación de video, es decir, anotación automática y recupero de los videos desde una base multimedia;
- Interacción humano computadora, es decir, reconocimiento gestual, seguimiento del iris ocular para ingresar datos a la computadora, etc.

- Monitoreo de tráfico, es decir, obtener en tiempo real las estadísticas de tráfico vehicular con el objetivo de redirigir el flujo de tráfico.
- Navegación vehicular, es decir, planificación del camino basado en video con capacidades de evitar los obstáculos.

Tracking puede ser definido como el problema de estimar la trayectoria y como se mueve un objeto en la imagen alrededor de una escena. Adicionalmente, un tracker puede proveer información del objeto como ser la orientación, área, forma, etc., de ese objeto. El tracking de objetos puede ser complejo debido a:

- Pérdida de información causada por la proyección del mundo 3D en una imagen 2D,
- Ruido en las imágenes,
- Complejidad en el movimiento del objeto,
- Oclusión total o parcial del objeto,
- Complejidad en las formas del objeto,
- Cambios en la iluminación de la escena, y
- Requerimientos de procesamiento en tiempo real.

Se puede simplificar el tracking mediante la imposición de restricciones en el movimiento y/o la apariencia de los objetos. Por ejemplo, se puede aplicar una restricción en el movimiento del objeto para que sea constante basado en información previa. En conclusión, información preliminar de los objetos, como el tamaño, la apariencia y la forma puede ser utilizado para simplificar el problema.

Se han propuesto diferentes soluciones para el tracking de objetos. Para llegar a una solución adecuada las diferentes propuestas se basan al menos en las siguientes preguntas:

- ¿Cuál representación del objeto es adecuado para tracking?
- ¿Cuáles características de la imagen debe ser utilizada?
- ¿Cómo debería ser modelado el movimiento, apariencia y forma del objeto?

Las respuestas a estas preguntas dependen del contexto/entorno en el que se realiza el seguimiento y el uso final para el que se solicita la información del tracking.

Como menciona [24] según la aplicación de RA que se desarrolle el tracking puede estar implementado en función de lo que se interesa conocer:

- La posición y orientación de la cámara en un sistema de coordenadas global

- La posición y orientación de la cámara en relación a un objeto de la escena real en cuestión
- La posición y orientación de la cámara en relación a la posición y orientación del cuadro de video anterior

3.7.1 Tracking mediante dispositivos físicos específicos

Los sensores proveen otra interesante oportunidad o soporte a los desarrolladores de aplicaciones de visión por computador. Muchos dispositivos cuentan con sensores tales como: acelerómetros, giróscopos, brújulas y GPS. Estos sensores son muy útiles a la hora de reconstruir la orientación de la cámara o la trayectoria en 3D.

El problema de extraer el movimiento de la cámara a partir de un conjunto de frames es desafiante, en términos de performance y precisión. Por ejemplo, SLAM (simultaneous localization and mapping), SfM (structure from motion) y otras soluciones pueden calcular la posición de la cámara e incluso la forma de los objetos que ve la cámara, pero estos métodos no son fáciles de implementar, calibrar y optimizar y además necesitan mucho poder de procesamiento.

Los sensores, no obstante, pueden proporcionar una estimación bastante precisa de la orientación del dispositivo. Para obtener resultados precisos el sensor debe ser utilizado sólo como punto de partida, para ser refinado mediante técnicas de visión por computador [37].

Por otro lado [38] detalla la fusión de tracking utilizando métodos visuales e inerciales y menciona como superar algunas limitaciones fundamentales en el tracking visual utilizando la combinación de estas dos técnicas.

- *Gran poder de cómputo*: el procesamiento de imagen usado para generar mediciones de la cámara es computacionalmente intensivo, sobre todo sin una secuencia de imagen continua y una predicción razonable de la pose de la cámara. Mediante la explotación de la información proporcionada por los sensores inerciales se puede obtener una predicción más exacta de la cámara. La predicción mejorada permite reducir el costo computacional del procesamiento de imágenes.
- *Movimientos lentos de la cámara*: los sistemas de tracking basados puramente en visión usualmente pueden manejar únicamente movimientos lentos de la cámara. Esto es debido

a la demanda computacional y también debido a las características inherentes de la cámara tales como el desenfoque de movimiento. Los sensores inerciales pueden ser muestreados a una tasa mucho más alta que las cámaras y los algoritmos son computacionalmente menos intensivos.

- *Oclusión*: si muchas características están ocluidas, la pose de la cámara no puede ser estimada utilizando técnicas de visión por computador. Los sensores inerciales, en general, brindan mediciones. Por lo tanto, la pose de la cámara puede ser actualizada cada tanto.
- *Jitter*: las técnicas de visión por computador tienden a introducir saltos (jitter). La razón es que la estimación actual de la pose por lo general se basa sólo en el conjunto de características visible. La cantidad de jitters aumenta cuando el conjunto de características cambia significativamente entre los frames.

3.7.2 Tracking basado en visión

Existen diferentes técnicas para realizar tracking basado en visión. Las mismas pueden categorizarse en dos grandes ramas:

- Tracking de marcadores
- Tracking basado en características naturales

Tracking de marcadores

Un marcador en el contexto de RA se refiere a lo que se conoce en idioma inglés como “fiducial markers”. Es tipo de marcador es cualquier objeto que se puede colocar en una escena para proporcionar un punto fijo de referencia. En RA, estos marcadores pueden proporcionar una interfaz entre el mundo físico y el contenido virtual que proporciona la RA, tales como modelos 3D o vídeos. Estos marcadores permiten al dispositivo que está generando el contenido virtual calcular la posición y la orientación, de su cámara. Cuando este cálculo se realiza en tiempo real, este proceso se conoce como tracking. El tracking de marcadores se puede lograr utilizando sólo información visual. En la práctica esto significa que el tracking de marcadores se puede lograr con sólo una cámara y una PC de escritorio; no es necesario ningún sensor adicional.

De acuerdo a [39] menciona que la información que proviene de la imagen de entrada se compone básicamente de características de la imagen que son fáciles de extraer de la escena. Las técnicas de tracking basadas en marcadores y las de características naturales (sin marcadores) se basan en este principio, con la diferencia de que el primero añade algunos marcadores o plantillas que originalmente no pertenecen a la escena con el fin de permitir el tracking.

En [24] se señala que el algoritmo de tracking basado en marcadores está formado por dos pasos principales:

- Detección e identificación de marcadores
- Estimación de la matriz de transformaciones geométricas de cada marcador detectado

En cada cuadro capturado se aplica un algoritmo de detección de marcadores seguido de una identificación de cada marcador.

Adicionalmente en [24] se detalla los tipos de marcadores utilizados: Template, ID-marker, DataMatrix y QR. Los marcadores ID-marker codifican un número de 9-bits (hasta 512 diferentes) en un patrón de 6 x 6, repitiendo los 9 bits 4 veces completando los 36 bits. Los marcadores DataMatrix y marcadores QR no fueron diseñados específicamente para RA, sino que su propósito inicial es codificar una serie de caracteres ASCII, como, por ejemplo, una dirección URL.

A continuación, se muestran los tipos de marcadores más utilizados:

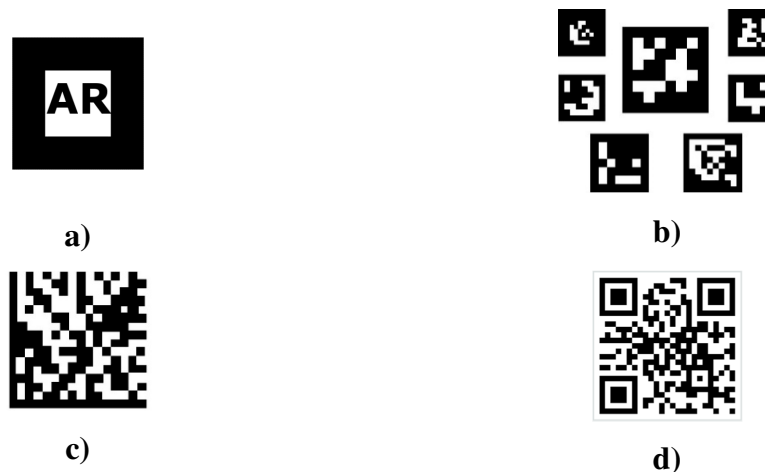


Figura 3.10 – Diferentes tipos de marcadores: a) Template; b) ID-marker; c) DataMatrix y d) QR (fuente: varios)

Tracking de características naturales

Las aplicaciones de visión por computador, incluido el reconocimiento de objetos, recuperación de imagen, odometría visual y RA, se basan, en gran medida, en el establecimiento de correspondencias entre puntos del mundo real similares que aparecen en las diversas imágenes. El método más aceptado y reconocido para lograr esto se basa en el reconocimiento de puntos claves que comprende dos etapas fundamentales:

- El primer paso corresponde a la detección de puntos clave (keypoints) en una imagen que contiene el objeto de interés.
- El segundo paso corresponde a la extracción de características cuya función es calcular una característica para describir la región que abarca el punto clave. Por lo tanto, se pueden encontrar correspondencias comparando las características de una imagen fuente y una imagen de consulta.

Ambos pasos han recibido una considerable cantidad de atención en las últimas décadas, con algoritmos que han sido diseñados para que sean más robustos a las variaciones minimizando el tiempo de procesamiento. En general, los descriptores de características pueden ser categorizados en dos clases:

- Basados en descriptor
- Basados en clasificación

Basados en Descriptor: se basa en el diseño de la característica sea robusta contra específicos tipos de deformaciones. Por lo general se centran en la escala o tamaño, rotación, corrección y normalización de la intensidad. Algunos enfoques basados en descriptores son: SIFT, GLOH y SURF. Las grandes necesidades computacionales de estos descriptores generalmente limitan su aplicabilidad a las aplicaciones en tiempo real. Los recientes intentos se han hecho para reducir el tiempo de procesamiento: ORB, BRISK, y FREAK.

Basado en Clasificación: a diferencia del anterior, requiere una etapa de entrenamiento off-line para aprender un conjunto de características provenientes de imágenes (del objeto de origen) lo que permite lograr insensibilidad a determinados tipos de deformaciones. Estos enfoques están diseñados para transferir gran parte de la carga computacional a la fase de entrenamiento con el fin de reducir el costo de comparación (matching) ejecutado de manera on-line al tiempo que aumenta su robustez.

Es bien sabido que lograr el reconocimiento de objetos invariantes ha sido uno de los mayores desafíos en visión por computador. Recientemente, ha habido un progreso significativo en el uso e implementación de algoritmos hacia la detección de características invariantes en imágenes más complejas.

Como menciona [40] es importante saber que los algoritmos basados en características han sido ampliamente utilizados como detectores de puntos debido a que las esquinas y bordes corresponden a las ubicaciones de imagen con alto contenido de información, lo que significa esto que pueden ser comparados entre las imágenes. Los detectores basados en características son precisos cuando los objetos a ser comparados tienen una esquina o borde distinguibles. En otras palabras, los detectores basados en características tienden a ser más adecuados sobre superficies planas y objetos dentro de una imagen dada. Además, los algoritmos basados en características no se desempeñan tan bien cuando las imágenes están sujetas a variaciones en la escala, la iluminación, la rotación o transformación afín. Para superar estas limitaciones, una nueva clase de algoritmo de comparación de imagen fue desarrollado simultáneamente. Estos algoritmos son conocidos como algoritmos basados en la textura debido a su capacidad para comparar características entre diferentes imágenes a pesar de la presencia de fondos texturizados y bordes bien definidos.

En [32] menciona que los trackers sin marcadores están basados en puntos de características naturales (frecuentemente llamado puntos de interés o puntos clave) visibles en el ambiente del usuario. A fin de obtener una precisión adecuada en la determinación de la pose, dichos puntos de características naturales deben reunir los siguientes requerimientos:

Procesamiento veloz

Debe ser posible para calcular grandes cantidades de características de los puntos y los descriptores asociados en tiempo real a fin de calcular la estimación de la pose en un tiempo aceptable.

Robustez con respecto a los cambios en las condiciones de luminosidad e imagen borrosa

El conjunto de características de puntos inclusive los descriptores calculados no deben variar significativamente bajo condiciones diferentes de luz o imagen borrosa. Ambos efectos son bastante comunes en particular en ambientes abiertos o exteriores.

Robustez frente a la observación de diferentes ángulos de visión.

Invariancia en la escala

En RA, los objetos que proporcionan los puntos característicos necesarios a menudo se observan desde diferentes distancias. La invariancia de escala se refiere al hecho de que los puntos característicos visibles desde una distancia bastante grande no desaparecerán.

Se debe considerar una buena característica local cuando cumple con las siguientes condiciones:

- *Localidad*: regiones pequeñas son menos sensibles a las deformaciones de la imagen deformaciones, y otras partes del objeto pueden ser ocluidas.
- *Pose invariancia*: El detector características de puntos puede seleccionar una posición canónica, la escala y la orientación para comparación subsecuente.
- *Distinción*: Los descriptores de características deberían permitir una alta tasa de detección (0,7 a 0,9) y una baja tasa de falsos positivos.
- *Repetitividad*: Debemos ser capaces de detectar los mismos puntos a pesar de los cambios en las condiciones de visión

El proceso de tracking en aplicaciones de RA consiste en comparar (matching) puntos 2D-3D y estimar la pose de la cámara. El matching determina una relación entre pares de puntos 2D-3D para establecer una restricción geométrica entre la imagen y los frames de referencia del objeto 3D. Diferentes algoritmos de estimación de la pose se desarrollaron, sin embargo, hay dos clases de algoritmos: analítico e iterativo. Los métodos analíticos son algoritmos que admiten un número finito de soluciones. Mientras que los métodos iterativos se basan en la minimización de un cierto criterio de error. La estimación de la pose de la cámara se basa en la extracción geométrica de primitivas lo que permite hacer el matching de los puntos 2D extraídos de la imagen con los puntos 3D conocidos del objeto. Para determinar la pose, los pares de puntos 2D-3D deben ser conocidos y la cámara se supone calibrada.

La calibración de la cámara determina el modelo geométrico de un objeto y el sistema de formación de la imagen que corresponde [41].

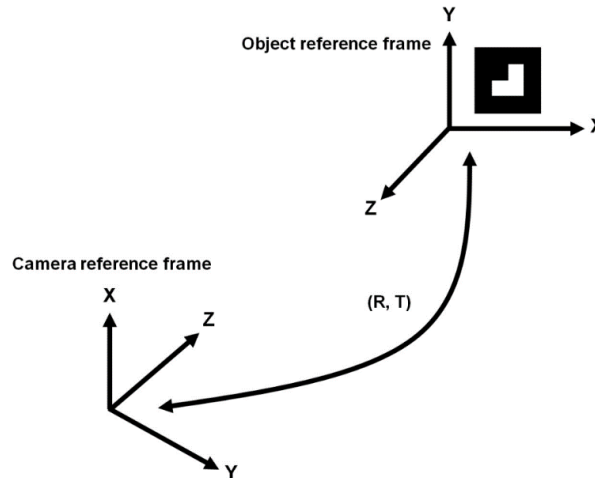


Figura 3.11 – Marco de referencia usado para estimación de la pose (fuente: Maidi, M. et al [41])

3.7.3 Tracking híbrido

Como menciona [42], los sistemas que utilizaban el tracking híbrido se hicieron disponibles durante la década de los noventa y utilizan por ejemplo compases electromagnéticos (magnetómetros), sensores de inclinación gravitatoria (inclinómetros) y giroscopios (mecánicos y ópticos) para el seguimiento de orientación y el seguimiento de posición ultrasónica, magnética y óptica, respectivamente. La utilización de técnicas de tracking híbridos son actualmente la manera más prometedora de lidiar con las dificultades que plantean los entornos de MRA en interiores y exteriores en general.

Por otro lado en [24] mencionan que en algunas aplicaciones de RA las técnicas de visión por computador no brindan una solución de tracking adecuada y por esta razón se han desarrollado métodos de tracking híbrido que consisten en combinar las salidas de dispositivos físicos y el análisis de video. Este tipo de tracking resulta efectivo para aplicaciones que requieren estimar la posición de la cámara con respecto a una escena estática, pero no se aplica al tracking de objetos en movimiento con una cámara estática.

De acuerdo a [43], las tecnologías de tracking pueden agruparse en tres categorías: *active-target*, *pasive-target* e inercial. Los sistemas de *active-target* incorporan emisores de señal y sensores puestos en un entorno preparado y calibrado. Ejemplos de tales sistemas utilizan señales magnéticas, ópticas, radio y acústicas.

Los sistemas *pasive-target* utilizan señales ambientales o de origen natural. Ejemplos incluyen brújulas que detectan el campo de la tierra y sistemas de visión que detectan marcadores fiduciaros intencionalmente colocados (por ejemplo, círculos, cuadrados) o características naturales. Los sistemas *inerciales* son aquellos que detectan fenómenos físicos creados por aceleración lineal y movimiento angular. Los sistemas híbridos intentan compensar las deficiencias de cada tecnología usando mediciones múltiples para producir resultados robustos combinando estas tecnologías con tracking basado en visión.

3.8 Reconocimiento de Objetos

De acuerdo a [19] algunos sistemas utilizan objetivos para ayudar en el tracking. Este tipo de tracking se conoce como tracking de landmark. Se conoce la posición Euclidiana de cada objetivo en el ambiente y esta información puede ser utilizada para inferir la posición de la cámara. Esta técnica requiere que estén visible dos o más características del objetivo proporcionando un registro preciso.

El número de características del objetivo depende del número de grados de libertad del punto de vista. El foco de los sistemas es determinar la posición de los objetos en la escena con respecto a la cámara.

Utilizar las características naturales del ambiente en lugar de objetivos elimina la restricción del movimiento de la cámara. Sin embargo, la detección de características naturales normalmente agrega suficiente complejidad computacional. En ambos casos, es decir, sistemas de tracking por características naturales o por objetivo, las características deben ser encontradas antes que sean rastreadas (tracked).

El enfoque basado en objetivos tiene ventajas y desventajas. Una desventaja es que el ambiente observado debe tener un mínimo número de objetivos sin obstrucción. También, la estabilidad de la pose disminuye con menos características visibles. La razón para el uso de las características naturales es eliminar el requisito de colocar objetivos en el medio ambiente. Para reconocer un objetivo es primordial obtener las características principales o atributos más relevantes de un objeto específico y compararlo contra una base de características de los objetos almacenados a fin de determinar la similitud alcanzada.

Como menciona [21] la extracción de puntos característicos es un paso usado frecuentemente para llegar a definir el objeto en sí dentro de una imagen. El objetivo principal es la obtención de los puntos característicos del objeto que nos permitan identificarlo nuevamente a partir de estos puntos.

Para la extracción de puntos característicos, existen una serie de pasos:

- *Detección*: consiste en encontrar el lugar de la imagen donde existen puntos característicos. La detección es el proceso en el cual se localizan las regiones de la imagen que contienen unas características determinadas, debido a su forma o textura.

Los retos actuales de los detectores son encontrar regiones características de distintos tamaños y con cualquier orientación. De esta forma se consigue ser invariante a tales efectos, lo cual permite detectar las mismas regiones en escenas donde varía el tamaño de los objetos y donde puedan estar orientados de cualquier forma. Asimismo es importante que la detección de puntos no se vea afectada por cambios de iluminación en la escena.

- *Descripción*: se realiza tras la detección. Una vez localizado el lugar, se procede a describir mediante una o más características la región situada alrededor del punto característico detectado. Estas características se almacenan frecuentemente en un vector para su uso posterior. Este proceso se lleva a cabo mediante el cómputo de la región característica que rodea a cada punto, concepto conocido como *descripción local*.
- *Comparación (Matching)*: para obtener resultados fiables en el proceso de comparación (matching) entre un par de imágenes es la tarea que algunos de los métodos de matching más populares están tratando de lograr. Pero ninguno de ellos ha sido universalmente aceptado. Es de destacar que la selección del método adecuado para completar una tarea de matching depende en gran medida del tipo de imagen y en las variaciones dentro de una imagen.

Algunos descriptores de características se han convertido en populares para el uso de navegación de robots en el interior y exterior de construcciones, tracking de objetos y reconocimiento, y algunas otras tareas que incluyen la comparación y registro de imágenes o secciones de imágenes. Es importante mencionar que los smartphones también utilizan reconocimiento de objetos en aplicaciones de, por ejemplo, RA, seguridad en basada en reconocimiento facial, e interfaces de usuario basados en reconocimiento gestual. Entre otros

métodos y por su precisión se podría utilizar el método SIFT como extractor de puntos de interés de los objetos y los describe de manera robusta contra cambios de traslación, escala y rotación. A este método se lo considera robusto con respecto a cambios geométrico gracias a su detector de región invariante y a su descriptor basado en la distribución de orientación. Sin embargo, SIFT consume mucho poder de procesamiento debido a la generación del descriptor y al proceso de matching. La detección e identificación de objetos sigue siendo un desafío en las aplicaciones de visión por computador debido al cambio de condiciones del ambiente.

Adicionalmente hay que considerar al reconocimiento de patrones como técnica para identificar y clasificar un objeto. En [21] se define al Reconocimiento de Patrones como la ciencia que se ocupa de los procesos sobre ingeniería, computación y matemáticas relacionados con objetos físicos o abstractos, con el propósito de extraer información que permita establecer propiedades de entre conjuntos de dichos objetos. Los patrones se obtienen a partir de los procesos de segmentación, extracción de características y descripción, donde cada objeto queda representado por una colección de descriptores. El sistema de reconocimiento debe asignar a cada objeto su categoría o clase (conjunto de entidades que comparten alguna característica que la diferencia del resto). Por lo tanto, una definición formal de Reconocimiento de objetos englobado dentro del reconocimiento de Patrones es la siguiente: es la categorización de datos de entrada en clases identificadas, por medio de la extracción de características significativas o atributos de los datos extraídos de un medio ambiente que contiene detalles irrelevantes.

En un problema de clasificación de patrones tradicional, se tienen varias clases que se caracterizan por sus estadísticas. Es decir, se asume que se ha observado en el pasado más de un ejemplo de cada clase y por lo tanto se puede clasificar cada clase por algunas propiedades del conjunto de observaciones pasadas (llamado un conjunto de entrenamiento). Un conjunto de entrenamiento puede estar caracterizado por un modelo paramétrico, en la que todo el conjunto está representado por un pequeño conjunto de parámetros tales como un vector de media y matriz de covarianza.

Descriptores de características

Uno de los objetivos propuestas en esta tesis, tal cual se mencionó en la sección *1.2 Objetivos*, es reconocer e identificar objetos, tales como construcciones edilicias y vehículos

de combate. Para lograr este objetivo una tarea importante es codificar la estructura local en un conjunto de puntos de una imagen con el fin de comparar entre las imágenes a pesar de los cambios visuales tales como cambios de escala, orientación, contraste, etc.

Los descriptores locales para cálculo de puntos de interés han demostrado ser muy exitosos en aplicaciones tales como reconocimiento de objetos, reconocimiento de texturas, ubicación de robots, minería de datos de video, reconocimiento de categorías de objetos, entre otros. Son distintivos, robustos a oclusión y no requieren segmentación. Recientes trabajos se han concentrado en hacer de estos descriptores invariantes a la transformación de imágenes. La idea es detectar áreas de imagen covariantes a una clase de transformaciones, que luego se utilizan como zonas de soporte para calcular descriptores invariantes [44].

Considerando a los detectores de regiones invariantes, la pregunta que resta hacer es cuál descriptor es el más apropiado para distinguir las regiones y si la selección del descriptor depende del detector de región. Hay un gran número de posibles descriptores y medidas de distancia asociados que hacen hincapié en diferentes propiedades de la imagen como intensidades de los píxeles, color, textura, bordes, etc.

Una característica local es un patrón en una imagen que difiere de su vecino inmediato. Está usualmente asociado con un cambio de la propiedad de la imagen o diversas propiedades simultáneamente, aunque no esté necesariamente localizado exactamente en este cambio. Las propiedades de la imagen comúnmente consideradas con la intensidad, color y la textura. Características locales pueden ser puntos, bordes o pequeñas porciones de una imagen. La invariancia en las características locales es una poderosa herramienta que ha sido aplicado exitosamente en un amplio rango de sistemas y aplicaciones [45].

Existen tres clases de categorías de detectores de características:

- 1) un tipo específico de características locales, que puede tener una interpretación semántica específica en el contexto de una determinada aplicación. Por ejemplo, bordes detectados en imágenes aéreas a menudo corresponden a las carreteras; detección de regiones se puede utilizar para identificar impurezas en algunas tareas de inspección; etc. Estas fueron las primeras aplicaciones para las que han sido propuestos detectores de características locales.
- 2) lo que las características representan realmente no es realmente relevante, siempre y cuando su ubicación puede determinarse con precisión y de manera estable en el tiempo.

Esto es, por ejemplo, la situación en la mayoría de las aplicaciones de matching o de tracking, y en especial para la calibración de la cámara o reconstrucción 3D.

- 3) un conjunto de características locales puede ser utilizado como una representación robusta de la imagen, permite reconocer objetos o escenas sin la necesidad de segmentar.

Las 3 categorías anteriores imponen sus propias limitaciones, y una buena característica de una aplicación puede ser inútil en el contexto de un problema diferente. Estas categorías pueden ser consideradas cuando se busca un detector de características adecuado para una aplicación determinada. Tradicionalmente, el término detector ha sido utilizado para referirse a la herramienta que extrae las características de una imagen, por ejemplo, una esquina, una región o un detector de bordes. La evaluación de los descriptores se ejecuta en el contexto de matching y reconocimiento de misma escena u objeto observado bajo diferentes condiciones de vista.

Muchas características son seleccionadas manualmente por el usuario dependiendo del dominio de aplicación. Sin embargo, el problema de la selección automática de características ha recibido atención significativa en la comunidad de reconocimiento de patrones. Los métodos de selección automática de características pueden dividirse en métodos de filtro y métodos de contenedor. El método de filtro intenta seleccionar las características basadas en un criterio general, por ejemplo, las características deben ser no correlacionadas [36].

Los métodos de contenedor seleccionan las características basadas en la utilidad de las características en un dominio específico del problema, por ejemplo, la performance en la clasificación utilizando un subconjunto de características.

PCA es un ejemplo de métodos de filtro para la reducción de características. PCA implica la transformación de un número de variables correlacionadas en un número (pequeño) de variables no correlacionadas llamada componentes principales. El primer componente principal representa la mayor cantidad de variabilidad en los datos como sea posible y cada componente subsiguiente representa como gran parte de la variabilidad restante como sea posible.

Un método de contenedor selecciona las características de manera discriminatoria para seguir (tracking) una clase particular de objetos es el caso del algoritmo Adaboost.

Adaboost es un método para encontrar un clasificador fuerte basado en una combinación de clasificadores débiles moderadamente inexactos. Teniendo en cuenta un amplio conjunto de características, un clasificador puede ser entrenado para cada característica. Adaboost, descubrirá una combinación ponderada de los clasificadores (características que representan) que maximizan el rendimiento de la clasificación del algoritmo.

3.9 Arquitectura de software para aplicaciones de RA

De acuerdo a [26] un sistema de RA requiere de un conjunto de componentes para cumplimentar con sus objetivos básicos: el reconocimiento en tiempo real del contexto del mundo físico, la registración de los objetos físicos de destino con sus objetos virtuales correspondientes, visualización del contenido de RA y el manejo de la interacción del usuario.

También, hay que considerar que la arquitectura global de una aplicación de RA es responsable por:

- Procesar el contenido incluyendo contenido adicional multimedia
- Procesar las entradas de los usuarios
- Procesar el contexto provisto por los sensores
- Administrar la presentación del resultado final (auditivo, visual, háptico, etc)
- Administrar la comunicación con los servicios adicionales

En cuanto a los procesos o tareas centrales que debe realizar una aplicación de RA se pueden ubicar la carga principal del lado del cliente, del lado del servidor o una distribución de carga equilibrada. Si consideramos una aplicación típica de RA que utiliza una sola fuente de vídeo, tanto para el tracking como para la proyección visual, el “pipeline” de procesamiento se debería componer de las siguientes tareas principales: adquisición de vídeo, tracking, cómputo de aplicaciones, renderizado y visualización. La descarga de algunas de estas tareas a un servidor permite ejecutar aplicaciones distribuidas, pero para ello se debe tener cautela en el uso del ancho de banda de red disponible. En [23] se detalla que existen 4 niveles diferentes de externalizar las tareas al servidor:

1. Toda la carga es procesada por el cliente independizándose del servidor y de la infraestructura que lo soporta.

2. Se externaliza al servidor únicamente la tarea de tracking. En este caso se enviará una imagen pre-procesada o un video comprimido para calcular la posición y orientación correspondiente. La ventaja de este tipo de externalización es que la tarea computacional es cargada al servidor mientras que los otros detalles de la aplicación son manejados por el cliente. Las dependencias entre el cliente y el servidor son mínimas. Por ejemplo, mientras que el tracking de un marcador puede ser ejecutado en tiempo real desde el cliente a través de un dispositivo móvil, el tracking de características naturales se puede beneficiar si se ejecuta desde el servidor por su gran capacidad de cómputo. No obstante, los procesadores de los nuevos dispositivos móviles son muy poderosos y esta tendencia podría cambiar drásticamente.
3. Se externaliza al servidor las tareas de tracking y de la aplicación. En este caso la aplicación envía al cliente un flujo de comandos gráficos o de información gráfica para que sea, posteriormente, renderizado el objeto virtual en el lado del cliente
4. Las tareas de tracking, el cómputo de la aplicación y el renderizado es enviado al servidor. Estas 3 tareas resultan las de mayor procesamiento de cómputo y el objetivo es enviar toda la carga al servidor.

En la siguiente figura se grafica los cuatro niveles de tareas para el procesamiento de RA entre el cliente y el servidor:

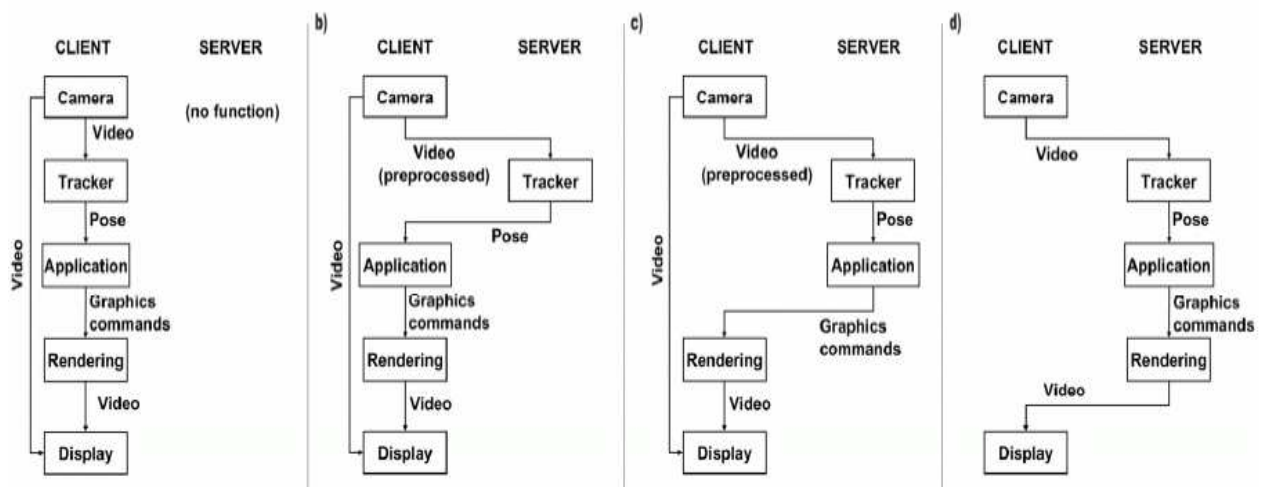


Figura 3.12 – Diferentes niveles de procesamiento de las tareas de RA. a) las tareas se ejecutan en el cliente; b) en el server se ejecuta la tarea de tracking; c) en el servidor se ejecutan las tareas de tracking y la lógica de la aplicación y d) todas las tareas se realizan en el servidor (fuente: Wagner, D. [23])

Una representación amplia de la arquitectura de los sistemas de RA lo aporta [46]. El autor identifica 6 formas diferentes de procesar la información en un contexto de RA:

- Reconocimiento local
- Reconocimiento y tracking local
- Reconocimiento remoto
- Registración local, reconocimiento remoto y tracking
- Tracking y registración local y reconocimiento remoto
- Tiempo real, registración remota, detección remota y presentación local

Reconocimiento local

El dispositivo detecta la presencia de la imagen objetivo (o los descriptores correspondientes) en un flujo de video.

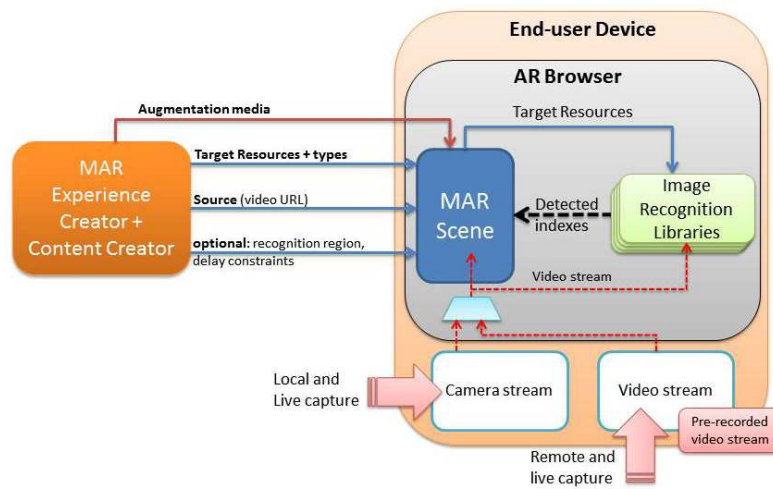


Figura 3.13 – Representación de la arquitectura de un sistema de RA procesando la información con Reconocimiento Local (fuente: Preda, M. [46])

Reconocimiento y tracking local

El dispositivo detecta la presencia de la imagen objetivo (o sus descriptores correspondientes) en un flujo de video, calcula la matriz de transformación (posición, orientación y escala) y aumenta con el correspondiente objeto gráfico.

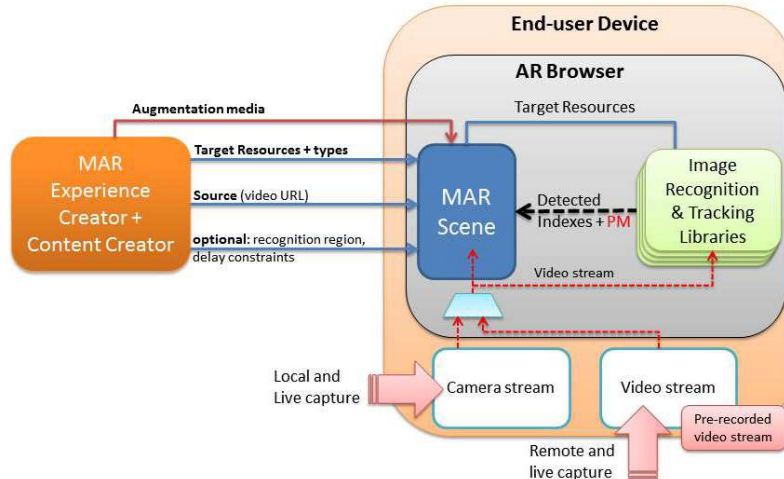


Figura 3.14 – Representación de la arquitectura de un sistema de RA procesando la información con Reconocimiento y Tracking Local (fuente: Preda, M. [46])

Reconocimiento remoto

El dispositivo envía la imagen objetivo (o sus descriptores correspondientes) a un servidor para que la procese. El servidor devuelve un identificador (ID) de la imagen detectada y procesada.

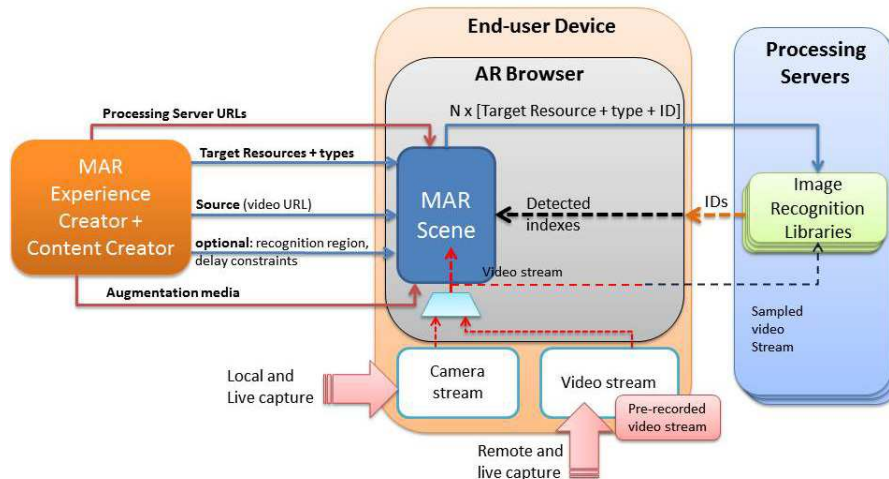


Figura 3.15 – Representación de la arquitectura de un sistema de RA procesando la información con Reconocimiento Remoto (fuente: Preda, M. [46])

Registración local, reconocimiento remoto y tracking

El dispositivo envía la imagen objetivo (o sus descriptores correspondientes) a un servidor remoto para que sea procesada, es decir, detectar la imagen para luego realizar el tracking

correspondiente. El servidor devuelve un identificador (ID) y la matriz de transformación calculada de los recursos detectados.

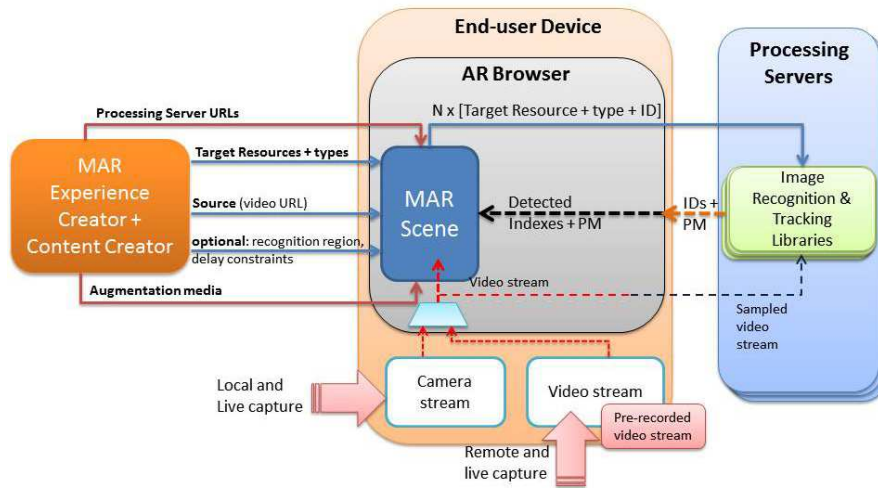


Figura 3.16 – Representación de la arquitectura de un sistema de RA procesando la información con Registro local, reconocimiento remoto y tracking (fuente: Preda, M. [46])

Tracking y registraci3n local y reconocimiento remoto

El dispositivo envía la imagen objetivo (o sus descriptores correspondientes) a un servidor remoto que se encarga de analizar los frames recibidos y detecta uno o más imágenes objetivas almacenadas en su base de datos local. El server devuelve la posici3n y tamaño de uno o más imágenes detectadas en el frame. Por medio de la posici3n y el tamaño, el dispositivo corta las imágenes objetivo del frame y las utiliza para el tracking local.

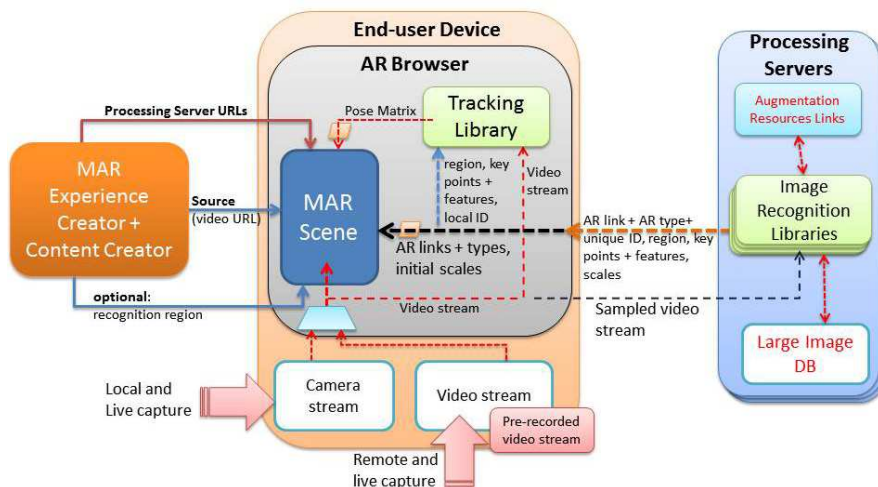


Figura 3.17 – Representaci3n de la arquitectura de un sistema de RA procesando la informaci3n con Tracking y registraci3n local y reconocimiento remoto (fuente: Preda, M. [46])

Tiempo real, registraci3n remota, detecci3n remota y presentaci3n local

El dispositivo envía las imágenes objetivo (o los descriptores correspondientes) a un servidor remoto para que detecte y realice el tracking de los recursos objetivo en el flujo de video. Adicionalmente, el servidor realiza la composici3n de aumentaci3n y renderizado, devolviendo al dispositivo la composici3n del flujo de video.

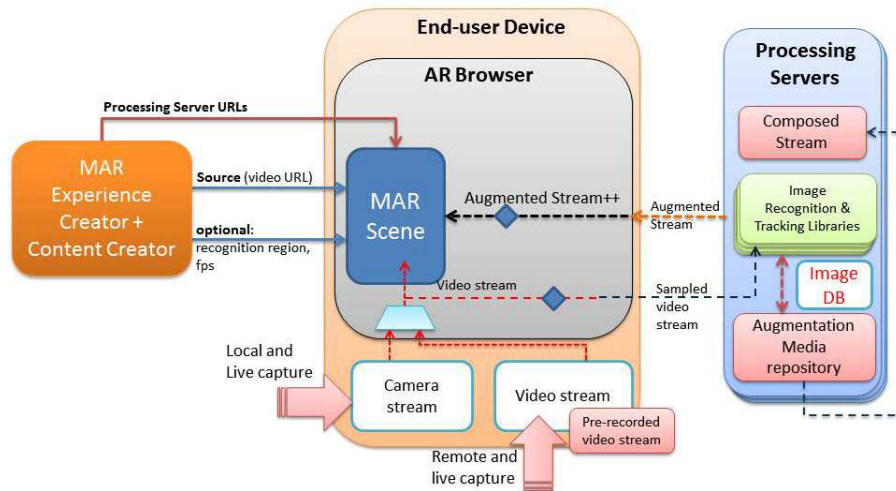


Figura 3.18 – Representaci3n de la arquitectura de un sistema de RA procesando la informaci3n con tiempo real, registraci3n remota, detecci3n remota y presentaci3n local
(fuente: Preda, M. [46])

3.10 Resumen del Capítulo

Este capítulo presentó una introducci3n a los aspectos técnicos relacionados con la RA. Se detalló las partes que forman parte de una aplicaci3n de RA, los métodos de entrada de datos en un sistema de RA, las técnicas y las tecnologías de visualizaci3n que propone la RA, se explicó qué es el tracking y las técnicas de reconocimiento de objetos basadas en detecci3n y extracci3n de características, culminando con un detalle de la arquitectura de software para aplicaciones de RA.

Cada uno de estos conceptos se detalla y se explican para entender cómo está formada una aplicaci3n de RA; sirviendo como base para comenzar a diseñar y desarrollar posteriormente el framework que se propone en el presente trabajo de tesis.

Capítulo 4

Conciencia Situacional y Realidad Aumentada

4.1 Introducción

En este Capítulo se detalla cómo se utiliza la RA para mejorar la CS de un combatiente en situaciones de estrés. Para ello se explica cómo se vinculan la CS y la RA en un campo de batalla para luego detallar y comparar un conjunto de proyectos militares en donde se utiliza la RA como tecnología principal para hacer frente a la CS.

4.2 La Conciencia Situacional y la Realidad Aumentada

En [1] el concepto de Soldado del Futuro fue desarrollado para identificar las capacidades de un soldado y llevarlas al campo de batalla. Es un concepto de cómo podría estar equipado el Soldado del Futuro. El Soldado del Futuro estará adaptado a cada área tecnológica teniendo especial énfasis en el rendimiento cognitivo para mejorar la eficacia del soldado y un aumento en el ritmo operativo.

Hay siete áreas principales aplicable al soldado:

- Rendimiento Humano y Entrenamiento
- Protección
- Letalidad
- Movilidad y Logística
- Redes
- Sensores

- Potencia y energía

Es importante resaltar el área de Rendimiento Humano ya que está relacionado con el desempeño de los soldados en el campo de batalla. Tener un correcto desempeño en una operación militar depende en gran medida en tener una correcta CS. La CS incluye la pérdida o percepción incompleta o cambio a elementos presentes en el entorno operacional del soldado. Este factor puede estar relacionado con las percepciones limitadas de un individuo de su entorno operacional, así como su capacidad de mantener un nivel alto de vigilancia de una serie de elementos.

Las futuras operaciones militares se apoyarán en herramientas de Comando, Control, Comunicaciones, Computación, Información, Inteligencia (C4I2) para un rendimiento óptimo en sus tareas asignadas en ambientes versátiles y hostiles. Los soldados portarán diminutas y potentes unidades centrales de procesamiento donde procesarán toda la información, capaz de controlar desde los parámetros fisiológicos, pasando por la gestión de las comunicaciones hasta la navegación o los datos de los subsistemas de armas y demás sensores asociados. En [47] se menciona que las soluciones críticas están centradas en mejorar:

- Comunicación
- Conciencia Situacional (CS)
- Comando y Control (C2)

Los ambientes virtuales y aumentados estarán en todas partes (ubicuidad) y apoyarán casi todas las facetas de la guerra incluyendo comunicaciones, visualización de datos, el control del sistema y capacitación.

El incremento en el uso de sensores de vehículos tanto en vehículos no tripulados como en lugares estratégicos se combinan para aumentar los requisitos de información a fin de proveer a los soldados datos que les sirva para cumplimentar su misión en el campo de batalla. Las unidades de combate se vuelven más dispersas y especializadas por lo que adquirir y mantener la Conciencia Situacional se vuelve más difícil. Por ejemplo, el aspecto predictivo de la CS se hace especialmente difícil en las operaciones urbanas, donde la línea de contacto visual con las fuerzas aliadas o amigas se hace difícil mantenerla durante un largo período de tiempo. En principio, algunas de estas dificultades se pueden superar mediante el

uso de una tecnología que puede organizar y presentar la información al usuario de forma automática. Un enfoque prometedor es la RA [48].

Por otro lado se puede mejorar la CS por medio del envío de información a través de los sensores dispersos geográficamente para luego visualizar dicha información utilizando como técnica la RA [49].

Es importante entender que la CS juega un papel crítico en las operaciones militares en cualquier terreno donde se desarrolle una misión. El comando y control dependen en gran medida de la CS, ya que las decisiones tienen que hacerse sobre la base de la situación actual.

La literatura apoya el potencial de la RA para mejorar la CS. Con el uso de la RA podría producir mejoras dramáticas en el rendimiento del soldado y proporcionar una gran ventaja en el combate.

4.3 Realidad Aumentada en Proyectos Militares

En lo que respecta al ámbito militar, recientemente se está utilizando la RA como soporte para mejorar la CS en la toma de decisiones en particular en operaciones militares. Muchas de las operaciones militares se desarrollan en entornos desconocidos. Estos complejos campos de batalla en 3D son muy exigentes e introducen muchos desafíos para el combatiente. Estos incluyen visibilidad limitada, falta de familiaridad con el medio ambiente, amenazas de francotirador, ocultamiento de fuerzas enemigas, mala comunicación y un problema de la localización e identificación de los enemigos y de las fuerzas aliadas. Para ello tener una CS amplia del terreno es vital para que la operación sea un éxito minimizando los efectos colaterales.

Livingston [6] [48] pasó a especificar diez observaciones importantes para el desarrollo de un exitoso sistema digital:

- No existen modelos maduros para la comunicación en ambientes urbanos
- Ningún modelo coherentemente integra la información del terreno, el clima, los océanos y el espacio para las operaciones urbanas conjuntas.
- Las unidades necesitan poder comunicarse entre sí y con unidades cercanas
- El seguimiento (tracking) de fuerzas aliadas en ambientes urbanos es importante

- Personalizar el cuadro de operaciones común (en inglés, common operating picture –COP-)
- La obtención de información debe ser automatizado de acuerdo a las prioridades y los roles del individuo.
- Las unidades deben actualizar los mapas y las bases de datos
- Existe la necesidad de información en 3D dentro de los edificios y bajo tierra
- La información debe estar rotulada
- Ayudar al usuario encontrar objetos de interés dentro del desorden visual

Una serie de programas de investigación han explorado los medios por los cuales la información de navegación y coordinación se pueden enviar al soldado utilizando técnicas de RA. Los programas del Soldado del Futuro han incursionado en la investigación y el desarrollo de componentes electrónicos aplicando la RA como tecnología para visualizar información del medio ambiente donde se produce el combate y de esta manera facilitar la toma de decisión ante situaciones hostiles.

4.3.1 Proyecto Eyekon

Hicks, Jeffrey et al [15] definen al proyecto EyeKon como un sistema de soporte a la toma de decisiones basado en agentes inteligentes instalados en una computadora portátil que transporta el soldado. El soldado visualiza información de objetivos y otra información en su armamento. El soldado no sólo ve objetos y posiciones sino también los medios para designar nuevos objetos como posibles amenazas y coordinar las acciones del equipo en respuesta a una situación que cambia dinámicamente. El proyecto tiene como objetivo desarrollar iconos inteligentes y descripciones que se superponen en el video del arma del soldado.

Los sistemas de armas deben ser optimizados no sólo para reducir al mínimo el peso total transportado por el soldado, sino también para reducir al mínimo la sobrecarga cognitiva. Demasiados datos en su pantalla y el soldado rápidamente se sienten abrumado con demasiada información afectando su conciencia situacional del medio que lo rodea.

Las funciones básicas se encuentran en una computadora portátil conectada, a través de una red inalámbrica segura, a una base de datos local y remota. Incorpora sensores que

brindan información en tiempo real (ej. Inercial, GPS, IR, etc). El software está compuesto por agentes que realizan consultas a la red a fin de monitorear información de amenazas y otros tipos de información estratégica para el soldado (ej. misión, estados, etc). Sobre la visión del arma se superimprime información utilizando técnicas de RA o a través del uso de unas gafas monocular.

EyeKon implica el diseño de una interfaz avanzada de usuario y es un subconjunto de RA de mejoras visuales aplicadas al concepto de Land Warrior.

Las características conceptuales son:

- Superponer iconos de situación en el armamento de un soldado
- Mejorar la conciencia situacional analizando el rango de amenazas
- Maximizar la capacidad de respuesta y minimizar los requisitos de peso y potencia

4.3.2 Proyecto BARS

El Naval Research Laboratory (NRL) desarrolló un sistema prototipo de realidad aumentada conocido como BARS (en inglés, Battlefield Augmented Reality System) [3]. Este sistema conecta a múltiples usuarios móviles junto con un centro de mando. BARS se centró en el desarrollo de un sistema digital para ayudar a resolver el creciente énfasis en las operaciones militares en terreno urbano. BARS realiza el seguimiento (tracking) de la posición y orientación de la cabeza del usuario y superpone gráficos junto con anotaciones que se alinean con los objetos reales en el campo visual del usuario. Varias unidades compartían una base de datos común, donde las personas podían optar por unirse a un canal determinado para acceder a los datos gráficos.

Una característica importante es que se utilizaron componentes comerciales con el objetivo de acceder a las mejoras o actualización de hardware apenas estuviesen disponibles en el mercado.

De acuerdo a [50] el proyecto se centró en la investigación de los aspectos relativos al Factor Humano e Interfaz de Usuario, haciendo principal énfasis en las siguientes características:

- *Visión “rayos X” y percepción de la profundidad:* una ventaja del proyecto era saber dónde se encontraban las tropas distribuidas en un terreno urbano; como también

establecer los puntos de interés en el entorno de los miembros del equipo. Ambos objetivos requieren que el sistema de Realidad Aumentada identifique objetos que se encuentran ocluidos a la visión del usuario. Para lograr se utilizó como principio la metáfora “visión de rayos X de superman”. La investigación se centró en el estudio de la profundidad de objetos gráficos.

- *Percepción Básica:* con el objetivo de mejorar la visión que se obtenía por medio del video provisto por las gafas, se investigó la posibilidad de perfeccionar la percepción del entorno. Para ello se analizaron aspectos como sensibilidad al contraste, percepción de los colores y percepción circundante.
- *Filtrado de la Información:* la sobre carga de información puede resultar un factor negativo a la hora de tomar una decisión. La cantidad de información puede llegar a ser demasiada para procesar en el ritmo dinámico de las operaciones militares, hasta el punto en que inhibe la capacidad del personal militar para completar las tareas asignadas. Se han desarrollado algoritmos para restringir la información que se muestra a los usuarios. El algoritmo de filtrado evolucionó a partir de un filtro basado en regiones a uno híbrido que se centra en el modelo de interacción espacial, un filtro basado en reglas y el concepto militar de un área de operaciones.
- *Selección de Objetos:* Para consultar, manipular, o actuar sobre los objetos, el usuario debe primero seleccionar estos objetos. El proyecto BARS permite a un usuario seleccionar objetos mediante la combinación de visión (utilizando el tracking de la cabeza) y señalización relativa dentro del campo visual utilizando tracking ocular o mouse 2D o 3D. La compleja naturaleza de la operación de selección hace que sea susceptible a errores. Para mitigar estos errores, se diseñó un algoritmo de selección probabilística multimodal (voz y gesto). Este algoritmo incorpora una jerarquía de objetos, diversos algoritmos de reconocimiento gestual y reconocimiento por voz.
- *Colaboración:* El sistema incluye una base de datos de información, que puede ser actualizado por otros usuarios. El intercambio de información a través del área de operación es un componente crítico del equipo para mejorar la Conciencia Situacional. Se diseñó un sistema de distribución de información de modo que estas actualizaciones se envían a través de la red de acuerdo con un esquema de prioridades. Por otro lado, se

utilizó Google Earth como API ya que ofrecía una adecuada visualización de edificios en 3D y, adicionalmente, permitió la rápida creación de prototipos.

- *Entrenamiento*: el entrenamiento requiere que los participantes se desempeñen en terrenos urbanos reales. El objetivo es realizar los entrenamientos en sitios reales con la superposición de objetos sintéticos (edificios, armamento, enemigos, etc). Un desafío en el entrenamiento utilizando Realidad Aumentada es que las fuerzas simuladas tienen que aparecer en la pantalla del usuario para dar la ilusión de que existen en el mundo real.

4.3.3 Proyecto iARM

En el año 2009, el Defense Advanced Research Projects Agency (DARPA) contrató a la empresa Tanagram Partners para desarrollar el proyecto Intelligent Augmented Reality Model (iARM) [51]. Básicamente, el objetivo de iARM era desarrollar un sistema digital integrado que podría mejorar significativamente la toma de decisiones del personal militar en entornos complejos a través de un sistema operativo integrado, un modelo de servicios de datos, y un HMD mejorado. El objetivo era que todos estos componentes trabajen juntos de una manera transparente permitiendo a los soldados percibir, comprender y, lo más importante, proyectar el mejor curso de acción para un mayor rendimiento para alcanzar los objetivos tácticos. El proyecto iARM abarcaba muchos de los atributos de la inteligencia artificial. En la siguiente figura se muestra el diseño conceptual del HMD y la visión del soldado a través de las gafas.

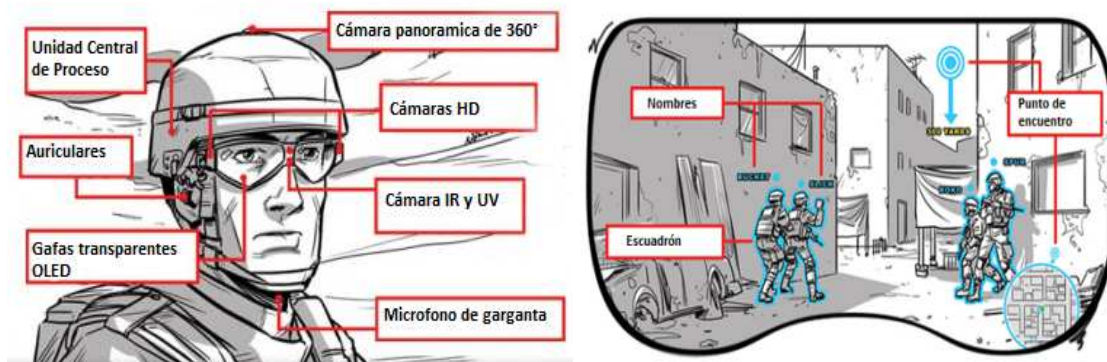


Figura 4.1 - Diseño conceptual del proyecto iARM
(fuente: Juhnke, J. et al [51])

4.3.4 Proyecto ULTRA-Vis

En [52] se detalla que el programa Urban Leader Tactical Response, Awareness and Visualization (ULTRA-Vis), soportado por el DARPA en su fase 1, ha desarrollado un prototipo de un sistema de RA para los soldados en el campo de batalla. El sistema ULTRA-Vis superpone iconografía gráfica a todo color en la escena local observada por el soldado. El programa desarrolló e integró un sistema de poco peso, una pantalla *see-through* holográfica de bajo consumo con un sistema de visión de tracking de posición y orientación. Usando el sistema ULTRA-Vis, un soldado puede visualizar la ubicación de otras fuerzas, los vehículos, los peligros y las aeronaves en el medio ambiente local, incluso cuando éstos no son visibles para el soldado. Además, el sistema puede ser utilizado para comunicar al soldado de una variedad de información tácticamente significativa incluyendo imágenes, rutas de navegación y alertas. El prototipo estará dotado para el reconocimiento gestual mediante el uso de un guante. Permitirá sobreimprimir símbolos en el campo de batalla en 3D, localizar objetivos enemigos y ubicar a las fuerzas aliadas. ULTRA-Vis provee a los escuadrones una ventaja táctica muy clara permitiendo la colaboración entre los integrantes del escuadrón. Posibilitará una alta conciencia situacional y la habilidad de tomar decisiones mientras se está en movimiento en el campo de operaciones.

En cuanto al equipamiento que forma parte del proyecto se destaca:

- Casco de combate con tecnología optical *see-through* con pantalla color
- Un unidad de procesamiento con arquitectura ARM

El sistema ULTRA-Vis incorpora un peso adicional adecuado al conjunto de componentes militares que lleva el combatiente. El prototipo procesa información obtenida de la red militar de datos obtenida por radio y sobreimprime información georegistrada por medio de iconos sobre el campo de visión del usuario.

En la siguiente figura se detalla los componentes del prototipo.



**Figura 4.2 – Componentes del prototipo ULTRA-Vis
(fuente: Roberts, R. et al [53])**

De acuerdo a [53] ULTRA-Vis posee una avanzada interfaz de usuario, que consta de las siguientes características:

- *Alertas operacionales*: informa al soldado sobre la inoperatividad de la radio; disturbios magnéticos que afectan a los componentes; GPS dañado, entre otros.
- *Sistema de información de estado de los componentes*: informa sobre el nivel de batería de la unidad de procesamiento; tiempo operacional; tiempo regional. El sistema de información es activado a demanda del usuario por medio de un botón que se encuentra en el casco.
- *Acceso a los menús del sistema*: los usuarios pueden acceder a los diferentes menús a fin de cambiar, por ejemplo, configuraciones del sistema.
- *Visualización de iconos de puntos de interés geo-referenciados*: los iconos o simbología militar son representados gráficamente dentro del campo de visión de la pantalla del usuario
- *Anillo de Conciencia Situacional*: es una herramienta gráfica intuitiva que permite al soldado tener un entendimiento en tiempo real de los aspectos circundantes en 360°. Estos aspectos están relacionados con el conocimiento de la ubicación de fuerzas aliadas, fuerzas enemigas y la ubicación de otros puntos de interés. Los iconos que se encuentran alrededor del anillo corresponden a la ubicación del soldado y la rotación del mismo.

En el siguiente gráfico se encuentra representada la interfaz de usuario del proyecto ULTRA-Vis

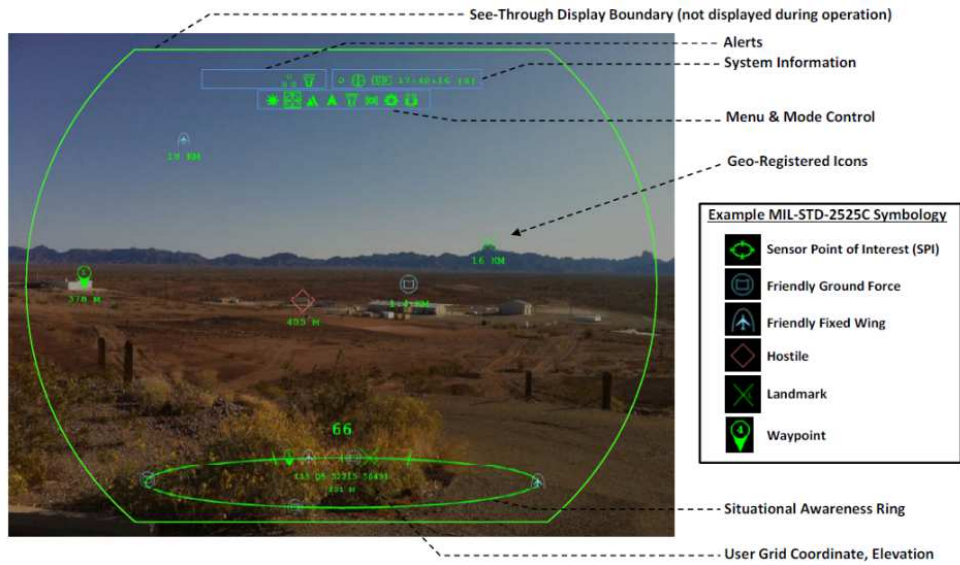


Figura 4.3 – Interfaz de usuario del prototipo ULTRA-Vis
(fuente: Roberts, R. et al [53])





Figura 4.4 – Información provista y componentes de hardware del proyecto ACR4 (fuente: ARA)

La evolución del proyecto ULTRA-Vis se denomina *ARC4*. Este proyecto, desarrollado por la empresa ARA¹, se caracteriza por mostrar íconos geo-registrados e información estratégica superponiéndose en el campo de visión del usuario. Se trata de un sistema de software que toma información provista por sensores externos, pudiéndose visualizar a través del HMD (Head-Mounted Display) del soldado. Una ventaja del software es que permite ejecutarse sobre dispositivos que soporten visión nocturna.

Las características principales del software ACR4 son las siguientes:

- *Visualización agnóstica*: permite renderizar información en cualquier dispositivo móvil de visión transparente –incluyendo dispositivos de visión nocturna-
- *Head Tracking*: implementa algoritmos de alta precisión para calcular la pose del usuario en ambientes externos
- *Visualización a través de visores (Heads-up)*: toda la información como iconos, anotaciones, etc., son visualizadas por medio de visores especiales.
- *Coordinación*: el software permite proyectar, entre los usuarios, la información como los iconos, las anotaciones y los mensajes en el campo de visión correspondiente.

¹ ARA: Applied Research Associates, Inc. <http://www.ara.com/>

4.4 Comparativa de tecnologías digitales militares que utilizan RA

Luego de haber analizado las tecnologías digitales que utilizan a la RA como soporte tecnológico para mejorar la CS en operaciones militares se construyó una tabla comparativa a fin de determinar la necesidad concreta de construir un framework propio de RA que sea aplicado en la construcción de aplicaciones y que sean ejecutables en tecnologías digitales provistas por el Ejército Argentino. Cabe destacar que los proyectos que se encuentran en las tablas comparativas forman parte de tecnologías integradas. El uso de RA para mejorar la CS, en alguno de los proyectos, es solo una pieza importante del conjunto.

A continuación se confeccionó una lista de elementos que hacen a la CS y que posteriormente se detallan como requisitos funcionales identificados y clasificados en la sección 6.3 *Identificación, selección y clasificación de requerimientos de la CS*. Estos elementos deben ser tenidos en cuenta en sistemas de RA para poder tener un mejor control y entendimiento del contexto del campo de operaciones con el objetivo de mejorar la CS en operaciones militares.

Estos elementos son los siguientes:

- *Reconocimiento del lugar de operaciones*: es de vital importancia tener una referencia precisa del lugar de operaciones. Conocer el terreno de operaciones aumenta la confianza y la percepción del entorno.
- *Identificación de infraestructura*: en un terreno hostil y poco conocido es de importancia identificar infraestructura con el objeto de utilizarlo ya sea como referencia de ubicación o para ejecutar una misión determinada.
- *Identificación de elementos del ambiente geográfico*: además de la identificación de infraestructura es importante obtener información contextual adicional para ello se debe identificar objetos, tales como referencias geográficas, piezas de artillería, u otro.
- *Alerta de amenazas mediante el uso de simbología descriptiva*: comprender el ambiente y sus amenazas circundantes le permite al operador tomar decisiones en el momento justo. Para ello se debe contar con una simbología militar, clara y precisa de las amenazas y alertas dentro de un área de acción establecida.

- *Localización de miembros aliados:* conocer la ubicación del resto de los operadores es de vital importancia desde un punto de vista estratégico y de seguridad para el conjunto de operadores
- *Tracking del trayecto de operaciones:* como método de reconocimiento del terreno y de puntos de chequeo operativos una opción es marcar en un mapa digital el trayecto que realizaron los operadores.
- *Comunicación entre el mando superior y el JPat:* la comunicación entre el mando central y el JPat forma parte de las tareas de comando y control. La comunicación debe ser hecha directamente con el jefe del grupo ya que será quién tomará las decisiones operativas principales.
- *Prioridad (filtrado) de información:* para no abrumar la capacidad de procesar la información contextual a la que debe someterse el operador, debe existir una primera capa de información, tal como, el radar e información contextual como las alertas. Las capas restantes deben visualizarse a pedido del operador a través de la introducción de gestos o de un comando vocal.

En la siguiente tabla se compara las funcionalidades implementadas en el resto de los frameworks militares:

Tabla 4.1 Elementos imprescindibles para mantener un alto nivel de CS

Elementos de la Conciencia Situacional (requerimientos funcionales)	EyeKon	BARS	iARM	ULTRA-Vis	RAIOM**
Reconocimiento del lugar de operaciones	X	X	X	X	X
Identificación de infraestructura			X	X	X
Identificación de elementos del ambiente geográfico	X		X	X	X
Alerta de amenazas mediante el uso de simbología descriptiva	X		X	X	X
Localización de miembros aliados			X	X	X
Tracking del trayecto de operaciones					X
Comunicación entre miembros y superiores		X	X	X	X
Prioridad (filtrado) de información	X	X	X	X	X

**** RAIOM: Realidad Aumentada para la Identificación de Objetivos Militares. Corresponde al framework propuesto de RA en el presente trabajo de tesis.**

A continuación, se listan las tecnologías que deben utilizarse para dar soporte técnico a los elementos que se detallaron precedentemente:

Tabla 4.2 Tecnologías utilizadas como soporte a la CS

Tecnología utilizada como soporte a la CS	EyeKon	BARS	iARM	ULTRA-Vis	RAIOM
Visualización mediante RA	X		X	X	X
Dispositivos liviano					X
Procesamiento distribuido					X
Sensores (GPS, acelerómetro, inerciales, etc)	X		X	X	X
Reconocimiento vocal		X	X	X	X
Reconocimiento gestual		X	X		X
Reconocimiento objetivos			X	X	X
Reconocimiento del lugar				X	X
Comunicación colaborativa		X	X	X	X
Software libre					X
Seguridad	X			X	X
Componentes COTS		X	X		X

4.5 Resumen del Capítulo

Este capítulo definió que es la CS y presentó un conjunto de proyectos militares que utilizan la RA como soporte tecnológico para mejorar la CS. De la investigación y análisis de los proyectos militares, encontramos que estos son insuficientes para cumplir con los requisitos específicos y la hipótesis de esta tesis. Analizando las dos tablas comparativas precedentes, es decir, por un lado, los elementos necesarios que forman parte de tener una CS amplia para tener el estado de situación de las amenazas circundantes y, por el otro lado, las tecnologías que dan soporte para que se cumpla lo antes detallado, se desprende que no todas las tecnologías cumplen con los requerimientos operacionales. Antes esta situación y siguiendo la metodología mencionada en la sección 1.4 - *Metodología de investigación*, se optó por desarrollar un framework de RA, denominado RAIOM (Realidad Aumentada para la Identificación de Amenazas y Objetivos Militares) centrado en resolver la problemática de la CS con el objetivo de poder desarrollar aplicaciones de RA que se ejecuten en una

arquitectura de sistemas acotada y sugerida en el presente trabajo de tesis. Cabe destacar que el componente digital más completo que se acercaba a lo solicitado por personal militar correspondía el proyecto ULTRA-Vis (ver sección 4.3.4 - *Proyecto ULTRA-Vis*), si bien es un componente que no es comercializable por el ejército de EEUU, sirvió como punto de análisis a fin de construir el proyecto de tesis presente.

Capítulo 5

Arquitectura y diseño de Software

5.1 Introducción

En este Capítulo se detalla la arquitectura de software propuesta. La importancia de diseñar un framework de uso específico. Adicionalmente se menciona la arquitectura en capas y componentes propuesto como también el ambiente de desarrollo y las propuestas de optimización a nivel software y hardware con el objetivo de mejorar la performance en el procesamiento de imágenes.

5.2 Framework y su importancia

La palabra inglesa "framework" (marco de trabajo, en español) define, en términos generales, un conjunto estandarizado de conceptos, prácticas y criterios para enfocar un tipo de problemática particular que sirve como referencia, para enfrentar y resolver nuevos problemas de índole similar².

Al centrarnos en las definiciones desde la perspectiva de desarrollo de software se puede afirmar que un framework es un diseño a gran escala que describe cómo una aplicación de software se descompone en un conjunto de objetos que interactúan y se relacionan. Generalmente está representado por un conjunto de clases abstractas cuyas instancias (objetos) interactúan entre sí. Una característica importante de un framework es el diseño de componentes que lo soporta. Adicionalmente se puede reutilizar la implementación, pero

² <http://es.wikipedia.org/wiki/Framework>

quedaría en un segundo plano ya que lo destacable de un framework es la reutilización de las funcionalidades y de los componentes que lo conforman.

Los frameworks se centran en respetar las características de los lenguajes orientados a objetos, es decir, abstracción de datos, polimorfismo y herencia. En este caso una clase abstracta representa una interfaz detrás de la cual las implementaciones pueden cambiar; el polimorfismo se refiere a la propiedad por la que es posible enviar mensajes sintácticamente iguales a objetos de tipos distintos mientras que la herencia es el mecanismo más utilizado para alcanzar la reutilización y la extensibilidad. Por medio de la herencia se pueden crear nuevas clases partiendo de una clase o de una jerarquía de clases preexistente evitando con ello el rediseño, la modificación y verificación de la parte ya implementada. La herencia facilita la creación de objetos a partir de otros ya existentes e implica que una subclase obtiene todo el comportamiento (métodos) y eventualmente los atributos (variables) de su superclase.

Un framework describe la arquitectura de un sistema orientado a objetos; las clases de objetos y cómo interactúan entre ellos. Describe como un software se descompone en objetos. Los frameworks se encuentran representados por un conjunto de clases, usualmente abstractas [54]. Adicionalmente [55] pone énfasis en definir a un framework como un conjunto de clases abstractas, especificando a una clase abstracta como una clase que no se instancia sino que es utilizada únicamente como una superclase. Debido que una clase abstracta no tiene instancias, se utiliza como un “modelo” para la creación de subclases en lugar de un “modelo” para la creación de objetos. Los frameworks los utilizan como diseños de sus componentes. Una manera en que una clase abstracta actúa como un diseño es que ésta especifica la interfaz de sus subclases. Cuando se indica que se desea utilizar una instancia de una clase abstracta, siempre significa una instancia de una clase que cumple con la interfaz de la clase abstracta. Adicionalmente menciona que los frameworks provén componentes reutilizables de software para el desarrollo de aplicaciones, integrando un conjunto de clases abstractas y definiendo una manera estándar de cómo las instancias de estas clases deben colaborar. Cabe destacar que el alcance de la reutilización en un framework puede ser significativamente mayor que el uso de librerías de funciones tradicionales o librerías convencionales de clases orientadas a objetos.

Generalmente, los frameworks se diferencian de las librerías de clases tradicionales ya que los primeros definen una aplicación “semi completa” que encierra estructuras de objetos

de un dominio específico mientras que las librerías de clases proporcionan una pequeña granularidad de reutilización. Por otro lado, los componentes que forman parte de un framework colaboran entre ellos a fin de proveer un modelo arquitectónico que sea familiar a un conjunto de aplicaciones específicas.

Es importante remarcar que un framework reduce la cantidad de código de una aplicación específica que pertenece a un dominio específico.

En la práctica, los frameworks y las librerías de clases son tecnologías complementarias. Los frameworks de manera frecuente utilizan librerías de clases internamente para simplificar el diseño del framework.

Por otro lado [56] resalta que la reutilización de arquitecturas software se define dentro un framework. En general, un framework se suele definir de la siguiente forma: *“Es un diseño reutilizable de todo o parte de un sistema, representado por un conjunto de clases abstractas y la forma en la cual sus instancias interactúan”*.

Otra forma de expresar este concepto es: *“Es el esqueleto de una aplicación que debe ser adaptado a necesidades concretas por el programador de la aplicación”*

En ambas definiciones podemos ver que un framework encapsula el patrón de la arquitectura software de un sistema o de alguna de sus partes, que puede además estar parcialmente instanciada.

Las principales ventajas que ofrecen los framework son la reducción del costo de los procesos de desarrollo de aplicaciones software para dominios específicos, y la mejora de la calidad del producto final.

En [57] define un framework de software como un esqueleto de programa que puede ser reutilizado o adaptado de acuerdo a reglas bien definidas para resolver un conjunto de problemas similares.

Los frameworks tienen como objetivo ofrecer una funcionalidad definida, auto contenido, siendo construidos usando patrones de diseño, y su característica principal es su alta cohesión y bajo acoplamiento. Para acceder a esa funcionalidad, se construyen piezas, objetos, llamados objetos calientes, que vinculan las necesidades del sistema con la funcionalidad que este presta. Esta funcionalidad, está constituida por objetos llamados fríos, que sufren poco o ningún cambio en la vida del framework, permitiendo la portabilidad entre distintos sistemas.

Por otro lado, es relevante destacar la importancia de la abstracción para construir un framework. Por último [57] detalla que la abstracción es un proceso conceptual por el cual un diseñador construye una vista simplificada de un sistema como un conjunto de interfaces. La implementación de estas interfaces en términos de entidades más detalladas se deja posteriormente para un mayor refinamiento. Un sistema complejo puede por lo tanto ser descrito en diferentes niveles de abstracción.

Uno de los objetivos principales en el diseño de un software es la reutilización del código y de las funcionalidades particulares para desarrollar una aplicación determinada.

Como detalla [54], una visión original en la reutilización de software se basó en la utilización de componentes. El interés comercial en el desarrollo orientado a objetos se centró también en la reutilización de código, de hecho, los diseños de software se focalizaron en esta idea dando lugar a los patrones de diseño. El concepto comenzó a madurarse en el tiempo dando lugar a la creación de estructuras de diseño más complejas que posteriormente se denominarían frameworks.

Pero los frameworks son formas intermedias, es decir, se reutiliza parte de código y parte del diseño. Lo destacado de un framework es que elimina la necesidad de realizar un nuevo diseño por completo utilizando como lenguaje de programación el paradigma orientado a objetos.

Cabe mencionar que la reutilización y los frameworks tienen muchas motivaciones. El principal de todas es la de optimizar y acelerar el tiempo de desarrollo que conlleva a ahorrar dinero. En el desarrollo de aplicaciones comerciales el tiempo de mercado (time to market) es de vital importancia estratégica y muchas compañías se encargan de desarrollar sus propios frameworks (o reutilizar algunos existentes) a fin de achicar el tiempo de salida al mercado de un producto de software específico.

Debemos considerar que el diseño de un framework se asemeja al diseño de la mayoría del software reutilizables. Generalmente se da comienzo con el análisis del dominio que se encarga, entre otras funciones, de obtener un número considerable de ejemplos. La primera versión del framework generalmente se diseña a fin de poder implementar los ejemplos, obteniéndose un framework de “caja blanca”. A partir de ese momento se puede utilizar el framework para construir aplicaciones de software. Sin embargo, estas aplicaciones son duras de desarrollar. La experiencia y el método de prueba y error, conduce a mejoras en el

diseño del framework pasando a ser un framework de “caja negra”. Este proceso puede continuar durante mucho tiempo hasta que, finalmente, el framework es lo suficientemente bueno donde las sugerencias de mejoras son raras. En este punto el diseñador tiene que decidir la finalización del framework para liberar una versión final [54].

5.3 Comparativa de frameworks de RA más utilizados

Conocidos también como SDK o simplemente frameworks, en [58] mencionan que estas herramientas están compuesta por varios componentes, tales como: reconocimiento, tracking y rendering, entre los principales. Teniendo en cuenta el ascenso en el uso de esta tecnología en diferentes segmentos, se concluye que existen diversos frameworks de RA en el mercado, sean estos de libre uso o comerciales. Cabe aclarar que existen en la actualidad una cantidad interesante de frameworks que ejecutan de manera óptima las características necesarias para desarrollar una aplicación de RA pero no son suficientes para cubrir los requerimientos funcionales necesarios para soportar la CS en operaciones militares.

Analizando el cuadro comparativo realizados en la sección 4.4 - *Comparativa de tecnologías digitales militares que utilizan RA* correspondiente a los framework de RA para el uso militar y teniendo en cuenta los requisitos funcionales que se identificaron y clasificaron en la sección 6.3 *Identificación, selección y clasificación de requerimientos de la CS*, en las siguientes tablas comparativas se representan los aspectos importantes para desarrollar aplicaciones completas de RA. Por otra parte, tomando como criterio de comparación lo aportado por [59], se definen un conjunto de parámetros funcionales tecnológicos entre los frameworks seleccionados:

- *Reconocimiento de objetos 3D*: reconocer formas tridimensionales de objetos. Los frameworks evaluados deben poder realizar tracking de objetos en 3D.
- *Reconocimiento de imágenes*: de la misma forma que el parámetro anterior, los frameworks seleccionados deben poder soportar el reconocimiento de imágenes en 2D.
- *Geolocalización*: se evaluó cuáles frameworks admiten el uso de geolocalización.
- *Reconocimiento sin uso de marcadores*: los frameworks seleccionados se evaluaron para identificar si soportan el uso de tecnología de RA que no requiera un marcador.

- *Reconocimiento de objetivos vía internet:* Se analizaron los frameworks hacen posible el reconocimiento de objetivos disponibles solo a través de internet en tiempo real.

Tabla 5.1 Parámetros funcionales tecnológicos

Parámetros funcionales tecnológicos	ARToolKit	Vuforia	EasyAR	Layar	DroidAR
Reconocimiento de objetos 3D	-	X	-	-	X
Reconocimiento de imágenes	X	X	X	X	X
Geolocalización	X	X	X	X	X
Reconocimiento sin uso de marcadores	-	X	-	-	-
Reconocimiento de objetivos vía internet	-	X	-	X	-

Adicionalmente a las funcionalidades antes mencionadas lo aportado por [60], ofrece una lista de parámetros comparativos centrándose en los aspectos no funcionales. Estos parámetros son los siguientes:

- *Disponibilidad off-line:* En el caso de operaciones militares en regiones donde la conexión es limitada, es necesario contar con aplicaciones que accedan a fuentes de datos almacenados en los dispositivos móviles.
- *Facilidad de integración con otras aplicaciones:* Aunque esto no es esencial, la integración con otros productos facilitaría el acceso o el intercambio de información. Por ejemplo, la aplicación desarrollada podría acceder directamente a la información de una base de datos y mostrar el resultado mediante el uso de RA.
- *Facilidad en ampliar el framework:* En general, no es fácil ampliar los frameworks ya que el código no está disponible y las restricciones de licencia prohíben extender el marco.
- *Tipo de licencia:* La mayoría de los frameworks están disponibles bajo una licencia propietaria, aunque algunos ofrecen una opción desarrollo libre (freeware) con funcionalidades limitadas y otros de código abierto (open source) que facilitan ampliar o modificara el framework.

- *Procesamiento distribuido*: La capacidad de procesar las imágenes en una arquitectura cliente servidor resulta de gran utilidad para computar grandes volúmenes de información.

Tabla 5.2 Parámetros no funcionales

Parámetros no funcionales	ARToolKit	Vuforia	EasyAR	Layar	DroidAR
Disponibilidad off-line	X	-	-	-	X
Facilidad de integración	-	-	-	-	-
Facilidad en extender el framework	X	-	-	-	X
Licencia	Open Source	Commercial	Freeware	Freeware	Open Source
Soporta procesamiento distribuido	-	-	-	-	-

Por último corresponde señalar que siguiendo lo propuesto por [60], en la evaluación se desestimaron los frameworks que:

- No soportaban GPS,
- No soportan dispositivos móviles,
- No soportaba Android como plataforma de desarrollo,
- No poseían soporte para sensores inerciales,
- Eran únicamente del tipo browser o visualizador 3D de RA,
- Estaban inactivos o eran considerados irrelevantes

5.4 Arquitectura de software por capas y componentes

En [61] definen el estilo en capas como una organización jerárquica tal que cada capa proporciona servicios a la capa inmediatamente superior y se sirve de las prestaciones que le brinda la inmediatamente inferior. En algunos ejemplares, las capas internas están ocultas a todas las demás, menos para las capas externas adyacentes, y excepto para funciones puntuales de exportación; en estos sistemas, los componentes implementan máquinas

virtuales en alguna de las capas de la jerarquía. En otros sistemas, las capas pueden ser sólo parcialmente opacas. En la práctica, las capas suelen ser entidades complejas, compuestas de varios paquetes o subsistemas. También mencionan que las ventajas del estilo en capas son obvias. Primero que nada, el estilo soporta un diseño basado en niveles de abstracción crecientes, lo cual a su vez permite a los implementadores la partición de un problema complejo en una secuencia de pasos incrementales. En segundo lugar, el estilo admite muy naturalmente optimizaciones y refinamientos. En tercer lugar, proporciona amplia reutilización.

Como menciona [57] un sistema de software se encuentra organizado por un conjunto de partes o componentes. El sistema en su conjunto, y cada uno de sus componentes, cumple una función que puede ser descrito como la prestación de un servicio. Con el fin de proporcionar su servicio, un componente por lo general se basa en los servicios prestados por otros componentes. Por el bien de la uniformidad, el sistema en su conjunto puede considerarse como un componente, que interactúa con un entorno definido externamente; el servicio provisto por el sistema se basa en suposiciones acerca de los servicios que el ambiente provee al sistema.

Un componente de software individual es un paquete de software, un servicio web, o un módulo que encapsula un conjunto de funciones relacionadas (o de datos) ³.

Todos los procesos del sistema son colocados en componentes separados de tal manera que todos los datos y funciones dentro de cada componente están semánticamente relacionados. Debido a este principio, con frecuencia se dice que los componentes son modulares y cohesivos.

Con respecto a la coordinación a lo largo del sistema, los componentes se comunican uno con el otro por medio de interfaces. Cuando un componente ofrece servicios al resto del sistema, éste adopta una interfaz proporcionada que especifica los servicios que otros componentes pueden utilizar, y cómo pueden hacerlo. Esta interfaz puede ser vista como una firma del componente - el cliente no necesita saber sobre los funcionamientos internos del

³ http://es.wikipedia.org/wiki/Ingenier%C3%ADa_de_software_basada_en_componentes

componente (su implementación) para hacer uso de ella. Este principio resulta en componentes referidos como encapsulados.

La reusabilidad es una importante característica de un componente de software de alta calidad. Los programadores deben diseñar e implementar componentes de software de una manera tal que diversos programas puedan reutilizarlos. Además, cuando los componentes de software interactúan directamente con los usuarios, debe ser considerada la prueba de usabilidad basada en componentes.

5.5 Ambiente de Desarrollo

El sistema operativo Android es un proyecto de código abierto iniciado por Google, Inc. Es una plataforma diseñada principalmente para dispositivos móviles con pantalla táctil, como teléfonos inteligentes o tabletas, y también para relojes inteligentes, televisores y automóviles. Android contiene una pila de software donde se incluye un sistema operativo, middleware y aplicaciones básicas para el usuario. Android es desarrollado por la Open Handset Alliance (OHA) un consorcio compuesto por fabricantes y desarrolladores de hardware, software y operadores de servicio., entre las que se encuentran Google, Inc., Motorola, HTC y Qualcomm. La arquitectura de Android está basada en el kernel de Linux versión 2.6. La plataforma de hardware principal de Android es la arquitectura ARM.

La estructura del sistema operativo Android se compone de aplicaciones que se ejecutan en un framework Java de aplicaciones orientadas a objetos sobre el núcleo de las bibliotecas de Java en una máquina virtual Dalvik con compilación en tiempo de ejecución. Las bibliotecas escritas en lenguaje C incluyen un administrador de interfaz gráfica (surface manager), un framework OpenCore, una base de datos relacional SQLite, una Interfaz de programación de API gráfica OpenGL ES 2.0 3D, un motor de renderizado WebKit, un motor gráfico SGL, SSL y una biblioteca estándar de C Bionic.

El siguiente gráfico muestra la arquitectura de Android. Como se puede ver está formada por cuatro capas. Una de las características más importantes es que todas las capas están basadas en software libre.

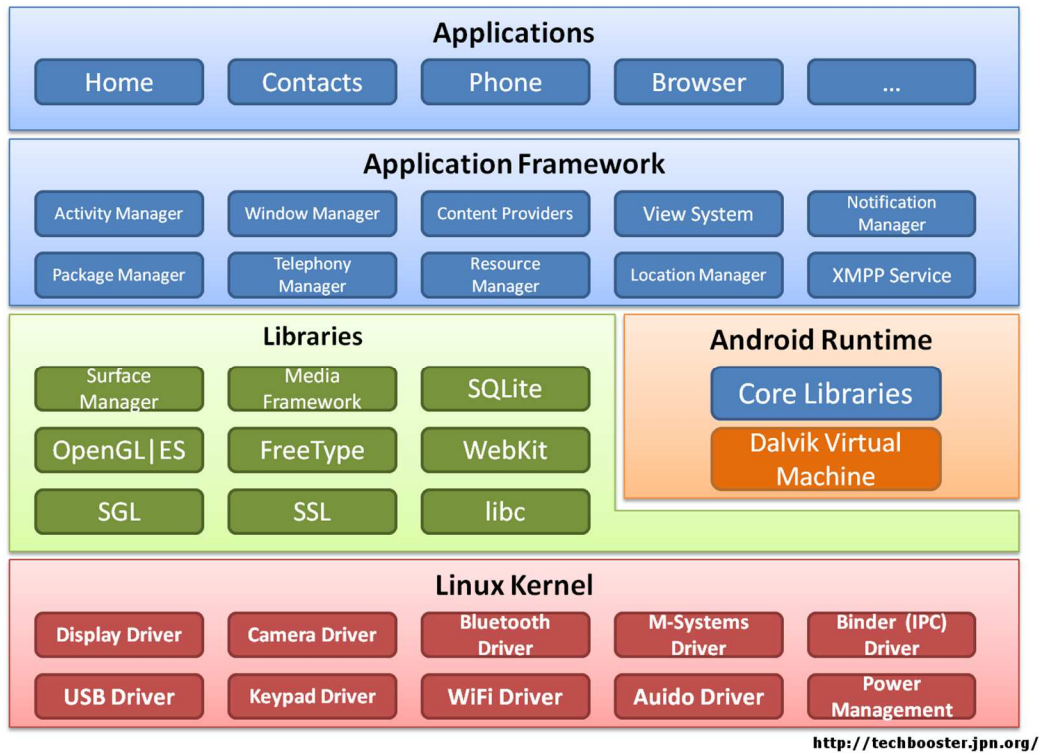


Figura 5.1 – Arquitectura de Android
(fuente: <http://techbooster.jpj.org/>)

Entonces, los componentes principales del sistema operativo de Android ⁴:

- *Aplicación:* un conjunto de aplicaciones núcleo están en esta capa de alto nivel. Incluye, cliente de email, aplicación de SMS, calendario, aplicación de mapas, web browser, etc. Todas estas aplicaciones están escritas en lenguaje Java.
- *Framework de Aplicación:* el framework de Aplicación es la base para el desarrollo de aplicaciones Android. Los principales componentes del Framework de Aplicación son: El Administrador de Actividades, el Administrador de Ventanas, Los Proveedores de Contenidos, el Sistema de Vistas, el Administrador de Notificaciones, el Administrador de Paquetes, el Administrador de Telefonía, el Administrador de Recursos, el Administrador de Locación y el Servicio de XMPP.
- *Librerías:* Android incluye un conjunto de librerías en C/C++ utilizados por varios componentes del sistema Android. Las librerías principales son: Framework de Medios,

⁴ <http://es.wikipedia.org/wiki/Android>

WebKit, SGL, OpenGL ES, FreeType, SQLite, etc. Los desarrolladores pueden acceder a estas librerías por medio del Framework de Aplicación.

- *Android Runtime*: cada Aplicación Android se ejecutan en su propio proceso otorgado por el SO y su propia instancia de la máquina virtual Dalvik. La máquina virtual Dalvik ha sido escrita para que un dispositivo pueda ejecutar múltiples máquinas virtuales eficientemente. La máquina virtual Dalvik ejecuta archivos de formato .dex (Dalvik Executable) el cual fue optimizado para utilizar mínimamente los recursos de CPU y memoria.

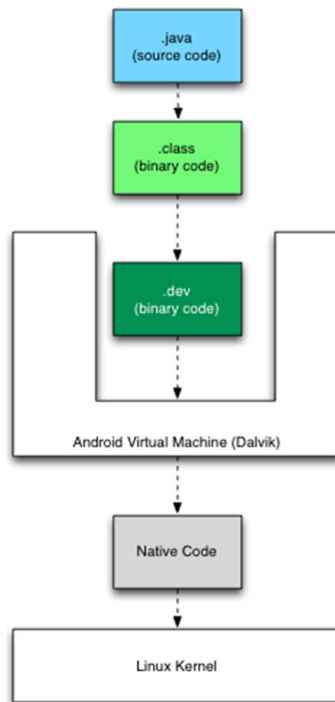


Figura 5.2 – Secuencia de ejecución de una aplicación Android
(fuente: <http://techbooster.jpj.org>)

- *Kernel de Linux*: Android está basado en Linux (Kernel versión 2.6) atendiendo servicios de sistemas tales como Administración de memoria, Administración de procesos, Pila de Red, Seguridad y Modelo de Controlador. El núcleo actúa como una capa de abstracción entre las aplicaciones y el hardware.

5.6 Técnicas de Optimización

Las aplicaciones de visión artificial en tiempo real que se ejecutan en dispositivos móviles o computadoras portátiles tienen una restricción importante en la cantidad de procesamiento que pueden realizar. Las técnicas de procesamiento de imágenes de RA, impulsan el desarrollo de plataformas integradas y la necesidad de procesadores más rápidos, es decir, más potentes a la vez. Las técnicas de visión por ordenador requieren una capacidad de cálculo intensivo. Las plataformas móviles de hoy tienen CPUs multi-core (Central Processing Unit) y GPUs integrados (Graphics Processing Units) para ofrecer el rendimiento exigido. El procesamiento paralelo en los sistemas multi-core ofrece una plataforma para alcanzar el rendimiento necesario en aplicaciones de visión por computador.

Muchos investigadores han estado explorando diferentes técnicas de optimización para mejorar el rendimiento de ejecución de aplicaciones de RA en los dispositivos móviles. Las técnicas de generación de código permiten escribir aplicaciones nativas y aprovechar múltiples núcleos, inclusive explotar las GPUs para los cálculos intensivos.

Por ejemplo, se comprobó que la programación en lenguaje nativo ofrece una mayor eficiencia comparado con la programación en Java. En [62] implementaron técnicas de multithreading a un sistema de detección de vasos sanguíneos desarrollado sobre sistema operativo Android, obteniendo un 92.77% de mejora en la performance comparado con una versión en Java en la misma plataforma de doble núcleo. Los algoritmos de visión por computador requieren mucho poder de procesamiento lo que resulta muy complicado que se ejecuten sobre dispositivos móviles. El procesamiento en paralelo sobre sistemas multi-core proveen una plataforma para aumentar el rendimiento que se requiere. En [63] han explotado computación en paralelo para usar las capacidades multi-core de un procesador Cortex-9 ARM para la optimización en tiempo real del proceso de decodificación de video en estéreo.

Estrategias diferentes fueron implementadas para mejorar los algoritmos de detección y extracción de características. En [64] se muestra una implementación GPGPU (General-Purpose Computing on Graphics Processing Units) del descriptor SURF llamado uSURF-ES, diseñado específicamente para dispositivos móviles. También existen diversos detectores de puntos diseñados para ejecutarse eficientemente sobre dispositivos móviles. Por ejemplo el algoritmo de detección FAST y sus variantes han sido probados exitosamente por muchas

aplicaciones móviles en tiempo real. FAST es eficiente y pueden encontrar puntos de interés con una precisión razonable pero no es invariante a la rotación y a la escala. Para hacer frente a estas limitaciones ORB, [65] utiliza FAST con un esquema piramidal de imagen para las operaciones de escala y orientación basado en momentos de intensidad. Aunque FAST y ORB son más rápidos que SURF, éstos son menos robustos con respecto a los cambios fotométricos y geométricos. Por otro lado, los frameworks de software integrados a las plataformas multi-core frecuentemente son necesarios para lograr un alto rendimiento en el sistema. En [66] se presenta un software basado en Android sobre una plataforma multi-core para dar soporte a un sistema multimedia en tiempo real.

5.6.1 Niveles de optimización

Las técnicas de visión por computador de RA requieren un alto procesamiento intensivo de cómputo. El objetivo es reducir la sobrecarga del procesador. Para superar esta situación, la mayoría de los dispositivos móviles se ejecutan en sistemas operativos con soporte para el multi-procesamiento, los cuales vienen con un conjunto de utilidades tales como cámaras, sensores, etc.

Del análisis de las técnicas de optimización surgen dos diferentes tipos de optimización que se agrupan a nivel software y a nivel hardware. Estos dos tipos de optimizaciones se complementan entre sí. Las optimizaciones a nivel de software se centralizan en la utilización de herramientas que utilizan técnicas de alto rendimiento para que los desarrolladores de aplicaciones puedan desarrollar óptimas aplicaciones de RA. Por otro lado, las optimizaciones a nivel de hardware están formadas por componentes de hardware de alto rendimiento con el fin de apoyar a las técnicas de optimización a nivel de software.

5.6.2 Optimización a nivel software

Dentro de las herramientas o procesos técnicos que forman parte de la optimización a nivel software se destaca el tipo de librerías que se utilizan para tal fin. Muchos de los frameworks que poseen como actividad central el procesamiento de imágenes utilizan, generalmente en su middleware o en la capa de librerías, una serie de librerías cuyo objetivo principal es la de mejorar el rendimiento general de las aplicaciones basadas en visión por computador. Estas

librerías tienen optimizados sus algoritmos para que puedan ejecutarse sobre una determinada arquitectura (por ejemplo, ARM con soporte multi-core y GPUs).

A continuación, se listan las librerías de mayor uso y soporte que existen en el mercado actualmente.

- *OpenCV*: es una librería de procesamiento de imágenes y visión por computador. Esta librería fue diseñada para ejecutarse de manera eficiente centrándose en la eficiencia de aplicaciones en tiempo real. OpenCV es multiplataforma, existiendo versiones para GNU/Linux, Mac OS X y Windows. Contiene más de 500 funciones que abarcan una gran gama de áreas en el proceso de visión, como reconocimiento de objetos (reconocimiento facial), calibración de cámaras, visión robótica, etc.
- *OpenGL ES*: OpenGL for Embedded Systems (ES), es una API para renderización gráfica para hardware acelerado ejecutándose sobre dispositivos móviles. OpenGL como tal no es compatible directamente con GPGPU o modelos de programación paralela en el sentido más estricto, no obstante se encontró que algunas de sus características para la representación de gráficos puede ser explotado para cálculos de datos en paralelo de la GPU ⁵.
- *Python Multiprocessing library*⁶: Multiprocessing es un paquete que permite crear nuevos procesos utilizando un API similar a la del módulo threading. Debido a que utiliza subprocesos en lugar de hilos (threads), permite llevar a cabo varias operaciones concurrentes sin las limitaciones del Global Interpreter Lock. Corre en sistemas Unix y Windows.
- *AsyncTask*⁷: Utilizado en la programación en Java en el SO Android. AsyncTask permite el uso adecuado y fácil del subproceso de interfaz de usuario. Esta clase le permite realizar operaciones en segundo plano y publicar resultados en el subproceso de interfaz de usuario sin tener que manipular subprocesos y / o controladores.

⁵ OpenGL ES, <https://www.khronos.org/opengles/>

⁶ Python Multiprocessing library: <https://docs.python.org/2/library/multiprocessing.html>

⁷ AsyncTask: <https://developer.android.com/reference/android/os/AsyncTask>

5.6.3 Optimización a nivel hardware

La optimización a nivel hardware se puede dividir en dos tipos pudiéndose complementar entre ellas: soporte de hardware para realizar multi-threading y multi-procesamiento y uso de la GPGPU. A continuación, se detallan brevemente estos dos tipos de optimizaciones aplicados al hardware ODROID-XU3 que se utilizó en el presente trabajo de tesis:

Soporte Multi-threading y Multi-processing

La velocidad y eficiencia de operaciones que requieren un proceso intensivo durante un tiempo prolongado generalmente se optimizan dividiendo en pequeñas operaciones que se ejecutan en múltiples hilos. En muchos dispositivos móviles que poseen CPUs con múltiples procesadores, el sistema puede ejecutar los hilos en paralelo en vez de ejecutar cada operación secuencialmente.

Como se mencionan en [67], los procesadores móviles modernos, que integran CPUs multi-core (múltiples núcleos), GPUs y otros aceleradores especiales, ofrecen múltiples opciones de portar aplicaciones de visión a diferentes dispositivos que soporten tecnologías multi-core.

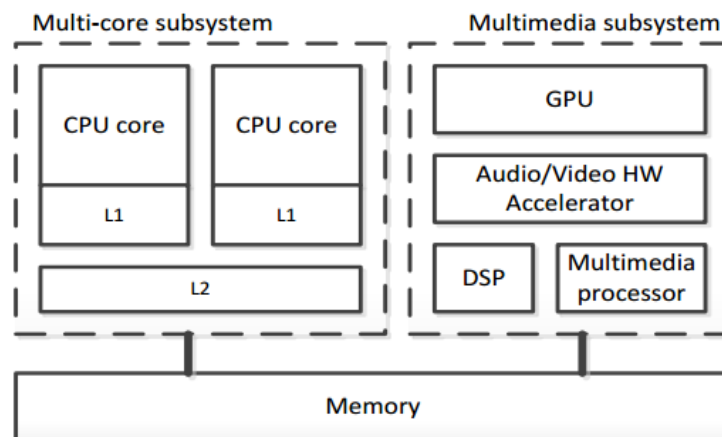


Figura 5.3 – Arquitectura común de un procesador con multi-core, GPU, DSP, etc., en un único chip
(fuente: Cheng, K. et al [67])

GPGPU

Los GPUs es un hardware de alto rendimiento que permite ejecutar de manera arbitraria cálculos paralelos por medio del GPGPU. Como la mayoría de los dispositivos móviles

modernos contienen una GPU es factible de aprovechar el poder de sus capacidades de procesamiento paralelo utilizando esta tecnología.

5.7 Resumen del Capítulo

Este capítulo presentó todo lo relativo a la construcción de un framework, su importancia para desarrollar aplicaciones de software e incluso se compararon algunos frameworks no militares de RA. Se consideró la arquitectura de capas y componentes a fin de facilitar la integración tecnológica y se mencionó algunas técnicas de optimización que fueron incorporadas en el framework con el objeto de ejecutar procesos asincrónicos. Lo importante a tener en cuenta es que si se toma como referencia principal la tabla comparativa detallada en la sección 4.4 - *Comparativa de tecnologías digitales militares que utilizan RA* y los requisitos funcionales detallados en la sección 6.3 - *Identificación, selección y clasificación de requerimientos de la CS* se llega a la conclusión que ninguno de los proyectos que incorporan la RA es completo en cuanto al aporte de funcionalidades que sirvan para hacer frente a la problemática central. Si bien los frameworks se acercan bastante a lo solicitado por los usuarios claves, el esfuerzo por completar las funcionalidades necesarias se consideró muy grande pues estos frameworks no fueron específicamente diseñados para el ámbito militar pues poseen particularidades que difieren bastante de aquellos frameworks utilizados para desarrollar aplicaciones en el ámbito civil. Cabe aclarar que los framework evaluados aportan parte de las funcionalidades e incluso la información como se debe presentar en los dispositivos de visualización, sean estos, smartphome, gafas de RA, etc,

Por lo tanto, en la siguiente Parte de la tesis se propone el framework que tiene como principal objetivo mejorar la CS en ambientes hostiles.

II. Framework Propuesto

Capítulo 6

Requerimientos iniciales

6.1 Introducción

Antes de detallar el diseño del framework propuesto se mencionará en esta sección como se arribó a la necesidad de construir un artefacto de software que sirviera para obtener información contextual de alertas y amenazas que mejorara la CS de los operadores. El objetivo inicial de esta parte del trabajo de tesis es identificar, clasificar y analizar los requisitos de los usuarios finales, establecer las características principales, sus prioridades y sus necesidades.

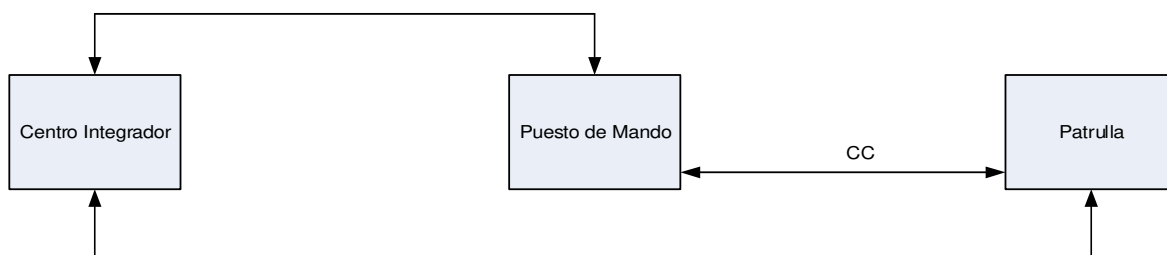
6.2 Esquema Conceptual de Operaciones

El Esquema Conceptual de Operaciones se refiere al ejercicio de la autoridad y la conducción y seguimiento por el mando operativo expresamente designado, sobre las fuerzas asignadas para el cumplimiento de la misión. Este esquema está representado por un flujo de información entre bloques que responden a los mandos y acciones a tomar en una operación militar. Cada bloque representa un puesto y cada uno tiene una función particular en el esquema propuesto. Los datos obtenidos del contexto, a través de los diferentes sensores (provistos por el Ejército y que actúan como proveedores de información contextual), son provistos por otro de los bloques que pertenece al esquema mencionado.

Por lo tanto, el esquema de mando y control tendrá tres bloques principales:

- *Centro Integrador (CII)*: se encarga de recibir información de todos los sensores distribuidos en el campo de operaciones. Procesa la información obtenida y la envía a los puestos de avanzada.

- *Puesto de Mando (JCa)*: Ejerce el comando y control de las tareas. Ordena y controla las actividades realizadas por las patrullas.
- *Patrulla (Pat)*: Ejecuta las tareas y las misiones encomendadas en el frente designado.
- *Jefe de la Patrulla (JPat)*: Será el principal operador que utilizará los componentes de hardware y de software propuestos en el presente trabajo. El JPat recibirá información de los otros bloques (también podrá enviar información a los otros bloques) y la representará visualmente mediante el uso de las gafas de RA.



**Figura 6.1 – Flujo de información entre bloques
(fuente: propia)**

Teniendo en cuenta los tres bloques definidos, se determinó el flujo de la información entre bloques con el objetivo de enviar datos procesados a los puestos de avanzada quienes serán los responsables de utilizar el equipamiento adecuado y la aplicación desarrollada extendida del framework propuesto.

A continuación, se detallada los lineamientos del flujo de información entre bloques:

- El flujo de la información se establece inicialmente entre el CII y el JCa. Luego JCa le envía información procesada a Pat. Por otro lado, el flujo de información puede ser directa entre el CII y la Pat que se encuentra en el frente de operaciones.
- El Comando y Control (CC) será únicamente entre el JCa y Pat.
- El bloque CII será el responsable de la carga de la información necesaria para la ejecución de la operación. También procesará la información obtenida del contexto en tiempo real.
- El bloque CII podrá realizar modificaciones a pedido del JCa y del operador de la Pat.
- La información de base estará almacenada en los dispositivos y el resto de la información adicional será a requerimiento de los usuarios. De necesitar una

información adicional en la ejecución de la operación el bloque CII será el responsable de proveerla.

- El bloque CII contará con un centro de Procesamiento de Datos cuyo fin es la de proveer a los otros bloques (JCa – Pat) información obtenida de los sensores externos.

Una vez que se definieron los bloques por donde fluirá la información, posteriormente, se determinó la ubicación esquemática de los integrantes, es decir, de los operadores principales que utilizarán el equipamiento correspondiente. Se acordó, en una primera instancia, que el equipamiento y la aplicación desarrollada serán utilizados por el JPat quién tendrá a su cargo la misión de visualizar la información contextual provista por los otros bloques.

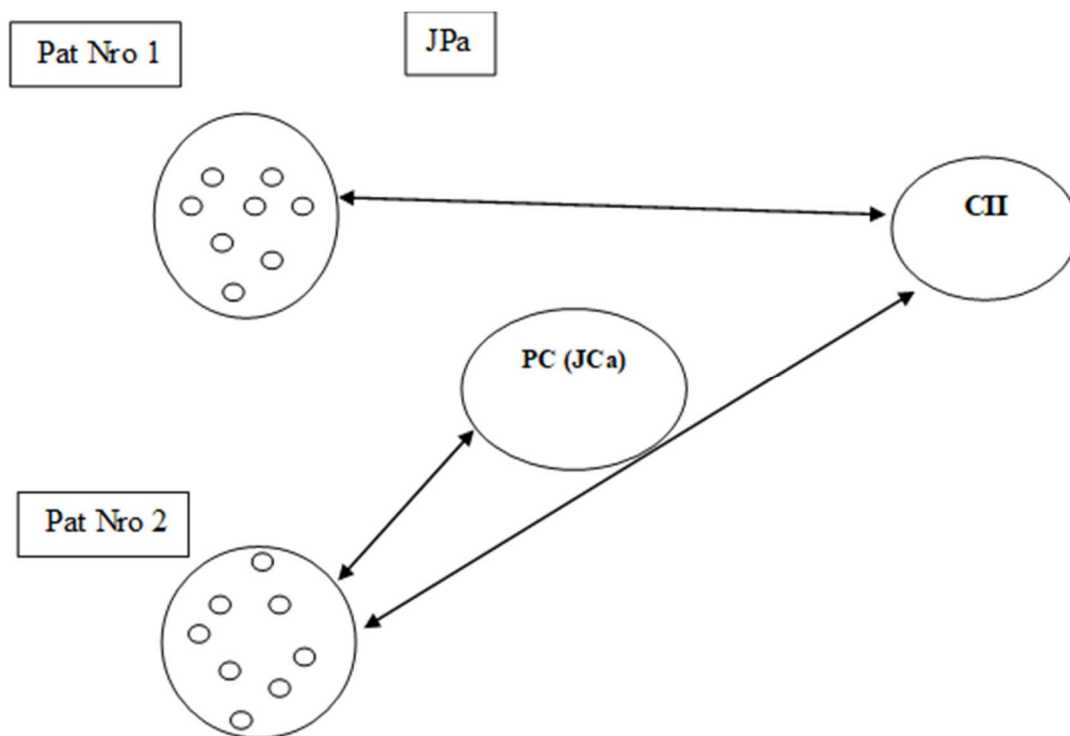


Figura 6.2 – Esquema de la ubicación de los integrantes/usuarios principales (fuente: propia)

Teniendo en cuenta el flujo de información de los bloques detallado en la figura 6.1 y siguiendo el esquema 6.2 de ubicación de los integrantes, la información que dispondrá cada bloque será la siguiente:

- El JPat dispondrá como parte de la información que visualizará los datos de orientación y posición del objetivo previamente establecido, como así también de cualquier otro dato de interés que hayan sido cargados por el CII.
- Las medidas de Coordinación y Control para la operación a ejecutar serán visualizadas por el operador y las mismas estarán establecidas por un Sistema de Información Geográfica (en inglés, Geographic Information System –GIS-).
- La información correspondiente al objetivo y al enemigo en general, será cargada por el CII en los dispositivos de procesamiento móvil que llevará el operador.

6.3 Identificación, selección y clasificación de requerimientos de la CS

Comprender los requisitos de los usuarios finales es una parte integral del diseño de sistemas de información y es fundamental para el éxito de los sistemas interactivos. Es importante mencionar que los sistemas y productos exitosos comienzan con una comprensión exhaustiva de las necesidades y requisitos de los usuarios.

Una de las características de la CS es la de no abrumar al operador/decisor con demasiada información pues la abundancia de ésta en vez de ser un factor útil se transforma en una sobrecarga cognitiva que afecta en la toma de decisiones.

Para hacer frente a las necesidades de los usuarios finales centrado en mejorar la CS es que se decidió identificar y analizar los requisitos funcionales aplicando una metodológica apropiada. Como se especifica en la norma ISO 9241-210 (2010), el DCU [68] comienza con una comprensión profunda de las necesidades y requisitos de los usuarios. Los beneficios pueden incluir una mayor productividad, una mejor calidad de trabajo, reducciones en los costos de soporte y capacitación, y una mayor satisfacción del usuario.

Para identificar las necesidades de los usuarios finales se utilizaron diferentes técnicas [69]:

Encuestas

Implican la administración de un conjunto de preguntas escritas a los usuarios finales. Se utilizó una técnica combinada de preguntas "cerradas" con respuestas fijas y preguntas

"abiertas", donde los encuestados fueron libres de responder como lo deseen. De esta manera se obtuvieron datos cuantitativos y cualitativos de los usuarios sobre las tareas existentes lo que permitió agrupar las principales preguntas en el nivel de *Percepción* del modelo de Endsley para luego entender los otros 2 niveles, es decir, *Compresión* y *Proyección*. Por otro lado, con el objetivo de determinar las condiciones presentes y futuras como parte del proceso de decisión de la CS dentro del espacio físico en donde se ejecutarán las operaciones militares, [70] aporta un conjunto de preguntas que forman parte de los elementos que responden a preguntas del tipo "quién, qué, dónde, cuándo y cómo", es decir, ayudan a determinar cuándo se trabaja con otras personas en un espacio físico compartido, se sabe con quién se está trabajando, qué están haciendo, dónde están trabajando, cuándo ocurren varios eventos y cómo ocurren esos eventos. Las siguientes tablas muestran estos elementos y se listan las preguntas que cada elemento puede responder. La Tabla 6.1 contiene aquellos elementos que se relacionan con el presente, y la Tabla 6.2 contiene aquellos que se relacionan con el pasado. Los elementos son todas cosas de sentido común que se ocupan de las interacciones entre una persona y el medio ambiente.

Tabla 6.1 Elementos de la conciencia del espacio de trabajo relacionados con el presente

Categoría	Elemento	Preguntas específicas
<i>Quién</i>	Presencia	¿Hay alguien en el área de trabajo?
	Identidad	¿Quién está participando? ¿Quién es ese?
	Autoría	¿Quién está haciendo eso?
<i>Qué</i>	Acción	¿Qué están haciendo?
	Intención	¿De qué objetivo es esa parte de acción?
	Artefacto	¿En qué objeto están trabajando?
<i>Dónde</i>	Ubicación	¿Dónde están trabajando?
	Contemplación	¿Dónde están mirando?
	Visión	¿Dónde pueden ver?
	Alcance	¿A dónde pueden llegar?

En [70] definen los elementos de la siguiente manera: la conciencia de la *presencia* y la *identidad* son simplemente el conocimiento de que hay otros individuos en el espacio de trabajo y determinan quiénes son, y la *autoría* implica el mapeo entre una acción y la persona que la lleva a cabo. La conciencia de las *acciones* y las *intenciones* es la comprensión de lo

que otra persona está haciendo, ya sea en detalle o en un nivel general. La conciencia del *artefacto* significa saber en qué objeto está trabajando una persona. La *ubicación*, la *contemplación* y la *visión* se relacionan con dónde está trabajando la persona, dónde está mirando y qué puede ver. La conciencia del *alcance* implica comprender el área del espacio de trabajo donde una persona puede cambiar las cosas, ya que a veces el alcance de una persona puede exceder su punto de vista.

Tabla 6.2 Elementos de la conciencia del espacio de trabajo relacionados con el pasado

Categoría	Elemento	Preguntas específicas
<i>Cómo</i>	Historial de acción	¿Cómo sucedió esa operación?
	Historial de artefacto	¿Cómo llegó a estar este artefacto en ese estado?
<i>Cuándo</i>	Historial de eventos	¿Cuándo sucedió ese evento?
<i>Quién (pasado)</i>	Historial de presencia	¿Quién estuvo aquí y cuándo?
<i>Dónde (pasado)</i>	Historial de ubicación	¿Dónde ha estado una persona?
<i>Qué (pasado)</i>	Historial de acción	¿Qué ha estado haciendo una persona?

Cabe aclarar que las encuestas son parte del trabajo de evaluación experimental que se encuentra en el Capítulo 9. Las mismas son realizadas a dos grupos de participantes con diferentes perfiles. También es importante aclarar que tanto la Tabla 6.1 y la Tabla 6.2 sirven como elementos esenciales para ir definiendo las preguntas finales que formarán parte del cuestionario que se utilizará para evaluar la CS de los participantes utilizando RAIOM.

Casos de Uso

El objetivo principal es proporcionar ejemplos de uso como una ayuda para comprender y aclarar los requisitos del usuario y para proporcionar una base para las pruebas de usabilidad posteriores.

Analizando la técnica de encuestas detallada anteriormente y de acuerdo a las preguntas que surgen del modelo mental del personal militar que ejecuta una operación en un espacio de trabajo hostil se arribaron a las siguientes preguntas para hacer frente a la CS:

- *¿Dónde estoy?:* se centra en reconocer el lugar geográfico donde se encuentra el operador realizando la misión.

- *¿Dónde están mis medios?:* se centra en reconocer dónde se encuentran los otros miembros de la Patrulla con respecto al J Pat.
- *Comando y Control:* se focaliza en reconocer a los otros miembros de la Patrulla, su ubicación y la relación con los centros de mando superiores.
- *¿Qué es eso?:* se focaliza en reconocer la infraestructura y elementos del ambiente geográfico que rodea al operador.
- *¿Dónde están las amenazas?:* se centra en determinar la posición geográfica de las potenciales amenazas.

En la siguiente figura se muestra como se agruparon las preguntas de acuerdo a los niveles del modelo propuesto por Endsley ampliado por las preguntas específicas de Gutwin et al. [70]. De esta manera se tiene una visión más concreta de la CS y posteriormente servirá para evaluar el diseño y usabilidad del producto software.

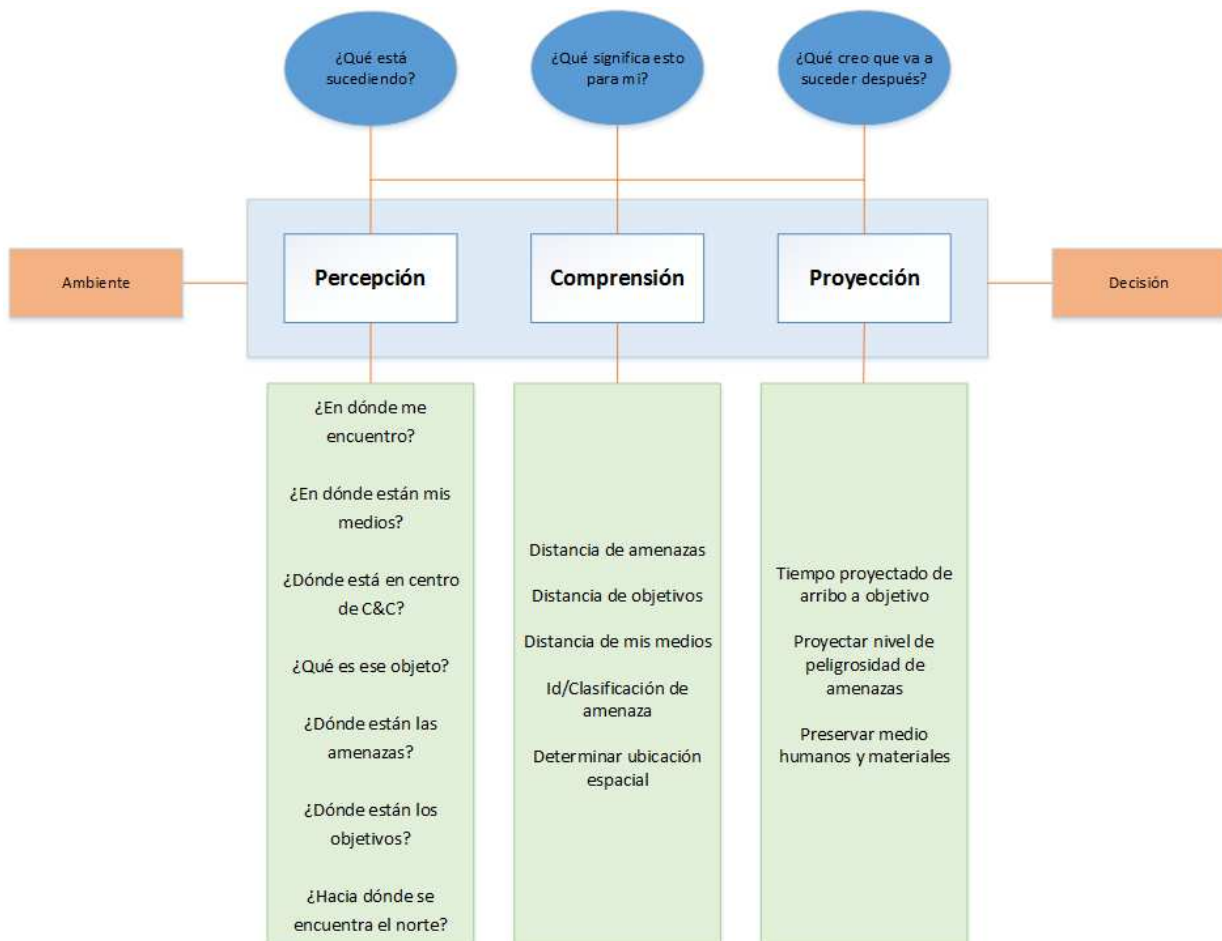


Figura 6.3 – Gráfico de las preguntas específicas de Gutwin et al. [70], en cada nivel de la CS del modelo de Endsley (fuente: propia)

En la siguiente tabla se detalla cómo se han mapeado [69] las preguntas de la encuesta donde se desprenden las tareas/funciones y se definen los requisitos funcionales de los usuarios finales con los CU correspondientes. En este contexto, los CU facilitan la interpretación de las necesidades del personal militar y suministran los medios necesarios para centrarse en el diseño de las interfaces de los despliegues –DCU- que se construirán más tarde.

Tabla 6.3 Tabla de definición de CU por niveles del modelo de Endsley

Niveles	Preguntas	Casos de Uso
Percepción	¿En dónde me encuentro?	Reconocer del lugar en dónde se realiza la operación Identificar de Infraestructura y elementos del Ambiente Geográfico.
	¿En dónde están mis medios?	Reconocer/Ubicar a los miembros de mi Unidad con respecto al Jefe y en el terreno.
	¿Dónde está el centro de C&C?	Ubicar a los elementos de mi Unidad y a los del escalón superior.
	¿Qué es ese objeto?	Identificar Infraestructura y elementos del Ambiente Geográfico
	¿Dónde están las amenazas?	Identificar las amenazas en el área de acción
	¿Dónde están los objetivos?	Identificar los objetivos en el área de acción
	¿Hacia dónde se encuentra el norte?	Identificar posición y orientación del operador
Comprensión	Distancia de amenazas	Calcular y ubicar amenazada en mapa
	Distancia de objetivos	Calcular y ubicar objetivos en mapa
	Distancia de mis medios	Calcular y ubicar medios en mapa
	Id/Clasificación de amenaza	Identificar objetos
	Determinar ubicación espacial	Calcular posición del operador
Proyección	Tiempo proyectado de arribo a objetivo	Visualizar Información contextual en el mapa
	Proyectar nivel de peligrosidad de amenazas	Visualizar Información contextual en el mapa

	Preservar medio humanos y materiales	Visualizar información del contexto y tomar decisiones
--	--------------------------------------	--

Tomando como referencia el cuadro anterior se ampliaron los CU detallando algunas características puntuales:

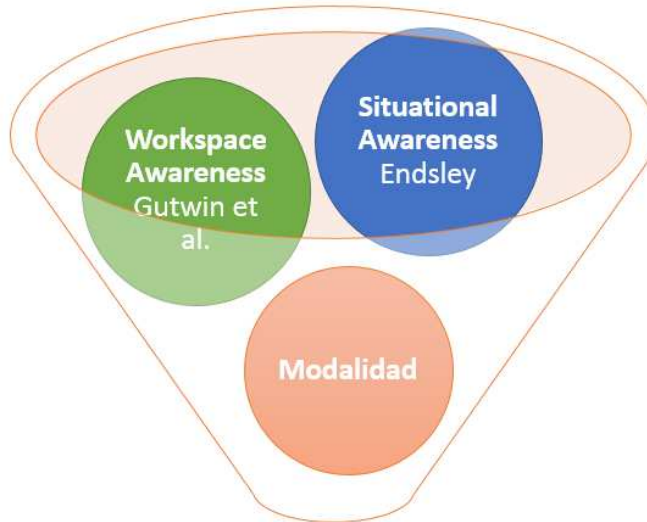
- *Interacción mediante reconocimiento gestual y vocal para la toma de datos:* los dispositivos de hardware deben ser livianos y el ingreso de datos al sistema debe hacerse de una manera rápida sin uso de periféricos externos tales como teclado y/o mouse.
- *Utilización de sensores (GPS, acelerómetros, cámaras, etc.):* utilizar los sensores provistos por los dispositivos para determinar la posición y orientación del operador principal del sistema.
- *Detección y reconocimiento de objetos tridimensionales:* en las operaciones militares en geografías poco conocidas es importante determinar puntos de referencias ya sea para ejecutar la operación o para determinar con mayor exactitud la posición de los operadores.
- *Identificación de los medios:* la identificación de los componentes aliados en una misión se convierte en una situación de alta prioridad estratégica ya que tanto el JPat como el JCa pueden saber la posición de los componentes (soldados) en el terreno.
- *Filtrado de información prioritaria:* la primera capa de información prioritaria corresponde a aquella referida a datos del contexto tales como alertas, amenazas, etc. Se detalla la información prioritaria de la capa inicial en la sección 7.3.1.5 - *Módulo Visualization*
- *Implementación e integración del prototipo en dispositivos móviles:* las aplicaciones que se desarrollen a partir del uso del framework propuesto deben poder ejecutarse en dispositivos móviles tales como gafas de RA, smartpone, tablets, etc.

Teniendo en cuenta lo anterior a continuación se detallan los requisitos no funcionales:

- Autónomo (poca dependencia de acceso a la red externa)
- *Omnidireccional:* debe permitir la comunicación entre los integrantes de la patrulla y el centro de Comando y Control

- *Liviano*: ingreso de datos o solicitud de información a través de reconocimiento gestual y vocal. Se deberá contar con componentes de hardware de bajo consumo.
- *Seguro*: deberá poseer cifrado del canal de información y de los datos que se almacenarán en el dispositivo móvil.
- *Código abierto*: el framework y las aplicaciones que se desarrollen deberán ejecutarse en sistemas operativos del tipo *open source* tales como Android.
- *Móvil*: las aplicaciones que se desarrollen a partir del framework deberán ejecutarse en dispositivos móviles tales como smartphones, tabletas y gafas de RA, entre las principales.

Luego de haber definido los CU gracias a la implementación del modelo de los tres niveles propuesto por Endsley y en las preguntas específicas que aportan Gutwin et al. [70], podemos combinar las propuestas antes mencionadas para la definición de un nuevo modelo para la identificación, selección y clasificación de los requisitos de la CS. Complementando las definiciones de CS aportadas en la sección 2.2 - *¿Qué es la CS?*, adicionalmente, identificamos los elementos de un entorno militar como lo sugieren Gutwin et al. [70], en su perfeccionamiento de la CS, en la Conciencia del Espacio de Trabajo (CET, en inglés, Workspace Awareness –WA-). Por ejemplo, en el contexto militar, la CS de RAIOM se refiere a la conciencia creada durante la interacción de varios individuos que pueden ser aliados o enemigos, en un espacio de trabajo compartido, limitado a los eventos, actividades y tareas que ocurren en ese espacio, ya sean asíncronos o sincrónico. Teniendo en cuenta todo lo anterior podemos decir que tres elementos caracterizan la CS (figura 6.4):



**Figura 6.4 – Combinación de propuestas para la definición de un nuevo modelo para la identificación, selección y clasificación de los requisitos de la CS.
(fuente: propia)**

Basándose en estos elementos y en la interacción con la información para facilitar la percepción, la comprensión y la proyección de la conciencia (Endsley), se propone un modelo tridimensional (figura 6.5) para gestionar con mayor precisión los requisitos de los usuarios finales que pretenden considerar y promover la conciencia. La primera dimensión, influenciada por la definición de Endsley, está determinada por la *Fase* o *Nivel* cognitivo que se pretenda lograr: *percepción, comprensión y/o proyección*; ya que no siempre será necesario lograr el máximo nivel de cognición de la situación. La segunda dimensión, sugerida por la definición de Gutwin y Greenberg, viene asociada a la *Característica*, es decir, a qué información o conjunto de informaciones se desea potenciar *quién, qué, dónde, cómo*, etc. En la tercera y última dimensión entra en juego la *Modalidad*, como por ejemplo, la RA y las diferentes posibilidades, no exclusivas, que se pueden poner en práctica para potenciar el feedback del sistema interactivo y la CS que éste feedback pretenden favorecer. Algunas de las interacciones consideradas son el *visual, el auditivo y/o el háptico*.

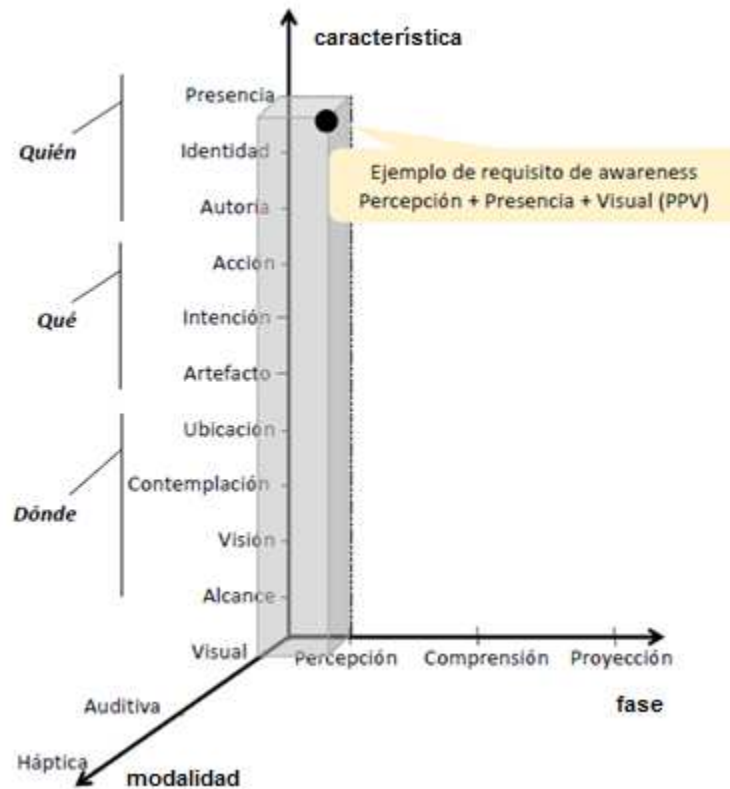


Figura 6.5 – Ejemplo de uso del modelo tridimensional propuesto para la identificación, selección y clasificación de requisitos de la CS (fuente: propia)

Un ejemplo del uso de este modelo de conciencia tridimensional puede ser el siguiente. Imagine una operación militar en la que se ha desplegado un sistema interactivo como RAIOM entre los diferentes actores involucrados. El actor debe ser consciente y, por lo tanto, percibir quiénes son sus aliados y sus enemigos. Este conocimiento se puede facilitar, por ejemplo, mediante el uso de una solución visual basada en símbolos militares. El sistema puede usar un símbolo militar de color azul para identificar aliados y un símbolo militar de color rojo para identificar enemigos. Debe poder entender dónde se encuentra (ubicación) y si la situación actual está en riesgo o no. Además, debe poder realizar cálculos y tomar decisiones en función de su situación y la de otros actores.

6.4 Resumen del Capítulo

En este Capítulo se analizó el Esquema Conceptual de Operaciones militares, se identificaron sus actores y sus funciones. Se centró en identificar y analizar los requisitos funcionales a fin de realizar el diseño de las aplicaciones de concepto y sus interfaces. Se utilizó un método que permita realizar un diseño centrado en los usuarios finales, para tal fin se optó por el método de DCU. Luego basándose en el modelo de Endsley de Percepción, Compresión y Proyección y en las preguntas específicas que aportan Gutwin et al. [70], se han mapeado las preguntas donde se desprenden las tareas y actividades realizadas por los usuarios con los CU respectivos. Ampliando los modelos mencionados anteriormente y sumándole el *ámbito* de aplicación se llega a la conclusión de que es posible diseñar un modelo tridimensional que sirva para la identificación, selección y clasificación de requisitos de la CS. El modelo tridimensional sugerido permite obtener de manera precisa los requisitos de los usuarios finales a través de la integración de los dos modelos antes mencionados, cada uno representado en un eje sobre el espacio más el eje que forma parte del *ámbito* de aplicación. Esta caracterización de la CS permite, no solo, organizar las diferentes definiciones previas de este concepto en una caracterización unificada, sino también incorporar la dimensión asociada con la solución identificada para facilitar la conciencia del usuario.

Una vez seleccionado el método para la identificación y análisis de requisitos de los usuarios y la identificación de los CU, en los capítulos siguientes se detalla la construcción del framework propuesto. Se realiza el diseño arquitectónico del sistema y del framework propuesto, se aborda la integración de técnicas de DCU con el prototipado de las aplicaciones de prueba que dará lugar a la realización de pruebas de concepto del uso del framework mediante una arquitectura Cliente / Servidor. Y por último se mide y se evalúa la calidad del producto.

Capítulo 7

Diseño de las Arquitecturas

7.1 Introducción

Debido a que el conjunto de requisitos de los usuarios identificados y clasificados en la sección 6.3 - *Identificación, selección y clasificación de requerimientos de la CS* pueden ser cambiantes, incompletos y dependientes de las capacidades de la tecnología que se pueden integrar bajo una arquitectura de sistema adecuada para afrontar las necesidades planteadas por los usuarios, se pensó en complementar las técnicas de DCU con un diseño ágil basado en SCRUM.

Se considera una arquitectura basada en *capas* para el diseño del middleware del framework propuesto pues facilita la integración tecnológica. Adicionalmente una de las características relevantes del framework propuesto radica en la arquitectura Cliente / Servidor, es decir, la capacidad de implementar (desplegar) las aplicaciones que se desarrollen en una arquitectura distribuida en donde en el lado del Servidor se ejecutarán las tareas más intensivas que corresponden al procesamiento de imágenes, tal cual se menciona en la sección 3.9 - *Arquitectura de software para aplicaciones de RA*.

Por último, se plantea como será la arquitectura y los componentes de hardware que forman parte del sistema en el cual se implementarán las soluciones de software que extiendan del framework de RA.

7.2 Metodología de desarrollo utilizada

Una de las características de la CS es la de no abrumar al operador/decisor con demasiada información pues la abundancia de ésta en vez de ser un factor útil se transforma en una sobrecarga cognitiva que afecta en la toma de decisiones. Para hacer frente a la problemática

de realizar un diseño aplicable a los usuarios finales es necesario desarrollar prototipos evolutivos e ir validando con los usuarios la aceptabilidad de uso del mismo.

Para el desarrollo del framework se evaluaron diferentes métodos, sobre todo aquellos que se ajustaban mejor al cambio constante por parte de los usuarios de las funcionalidades ya que al tratarse de la construcción de un marco de trabajo y de aplicaciones de prueba nuevas sin tener un conocimiento previo de la tecnología propuesta se hacía difícil interpretar tanto las necesidades como lo que deseaba experimentar el usuario. Para llevar a cabo el proceso de construcción del framework se basó en el desarrollo de pruebas de concepto individuales a modo de prototipos con el fin de que se demuestre el funcionamiento y el cumplimiento de lo solicitado. Para ello el diseño del framework y el desarrollo de las aplicaciones se basaron en una metodología combinada entre el uso de un método ágil y las técnicas de DCU mencionadas en la sección 6.3- *Identificación, selección y clasificación de requerimientos de la CS*. Por otro lado, se escogió este método de trabajo debido a que el segmento militar era poco explorado y era necesario diseñar, programar y mostrar a los usuarios claves el funcionamiento de cada parte del framework como unidades independientes, que posteriormente, se integrarían en el framework final.

Como menciona [71] el interés en la integración de DCU y los enfoques de desarrollo ágil está creciendo. Aunque las metodologías ágiles y DCU son dos enfoques de desarrollo de software diferentes, comparten principios comunes, tales como: diseño iterativo, participación del usuario, pruebas continuas y creación de prototipos. DCU es un proceso de diseño UX para interfaces de usuario (software) que se puede ejecutar de forma diferente. Sin embargo, es fundamental para esto la alineación con las necesidades del usuario y el procedimiento iterativo (incluso utilizando un método ágil), con pasos de proceso repetitivos.

Por las razones mencionadas anteriormente se optó por utilizar SCRUM⁸ como método ágil de desarrollo de software en conjunto con los principios de diseño que aporta el DCU resultando una metodología combinada SCRUM-DCU. Siguiendo el criterio de SCRUM-DCU, una de las principales tareas es la realización de prototipos evolutivos. Desarrollar prototipos de conceptos de diseño permite entender un diseño particular, un

⁸ SCRUM: <https://www.scrum.org/>

entendimiento en común de lo que el sistema podría hacer y cómo podría verse. Estos prototipos proporcionan la base para detectar problemas de diseño y deficiencias al principio del proceso de diseño. Sin el desarrollo de estos prototipos y modelos, se detectarán muchos problemas y deficiencias importantes en etapas tardías del proceso de diseño y desarrollo de la interfaz.

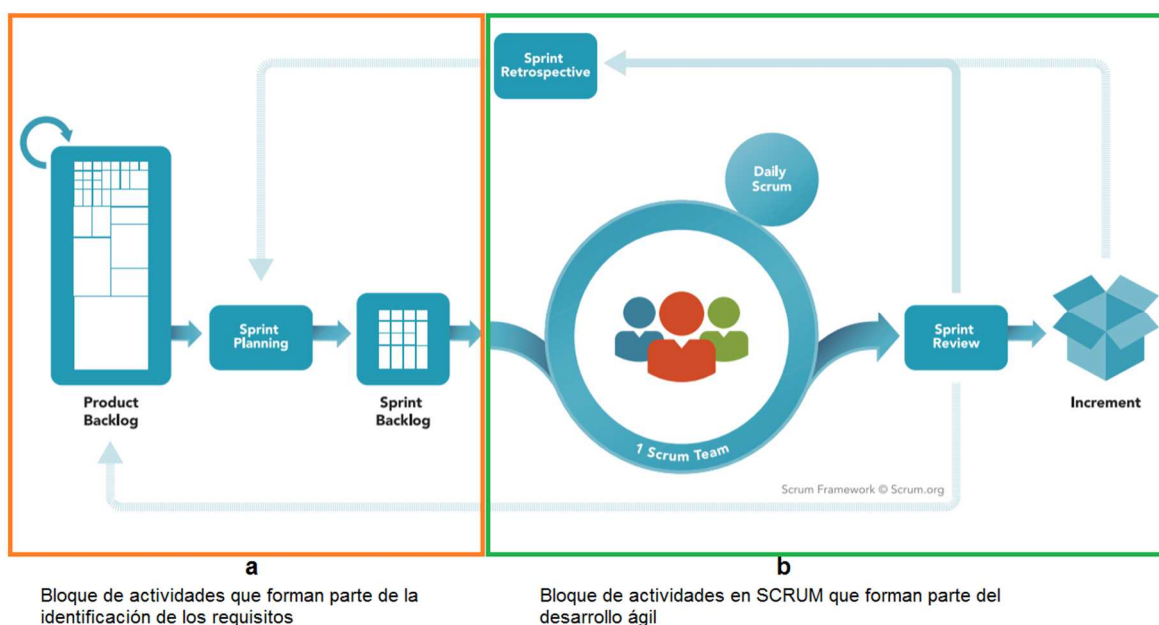


Figura 7.1 – SCRUM + DCU
(fuente: scrum.org)

En la figura anterior se encuentra cada tarea del proceso SCRUM, incluido los principios de diseño del DCU. El bloque de actividades a), que forman parte de la identificación de los requisitos, se hace énfasis especialmente a las preguntas de Gutmin at al. [70] (*Quién, Qué y Dónde* –ver sección 6.3 - *Identificación, selección y clasificación de requerimientos de la CS*). Por otro lado, DCU aporta técnicas de identificación de requisitos, tales como: entrevistas con usuarios principales (militares), identificación de tareas relevantes para los usuarios (considerando las preguntas *Quién, Qué y Dónde*), realización de actividades de prototipado, evaluación e iteración. Mientras que en b) se mencionan el conjunto de actividades en SCRUM que forman parte del desarrollo ágil priorizado por intereses de los usuarios.

A continuación, en la siguiente tabla, se resume las ventajas que aportan cada una de las metodologías utilizadas y la combinación de ellas en el proceso de identificación y

clasificación de requisitos y en diseño de los prototipos evolutivos enfocados en la CS de los usuarios:

Tabla 7.1 Ventajas en la combinación de las metodologías SCRUM y DCU (SCRUM+DCU)

Metodología	Ventajas
DCU	entrevistas con usuarios principales (militares), identificación de tareas relevantes para los usuarios (considerando las preguntas <i>Quién, Qué y Dónde</i>), realización de actividades de prototipado, evaluación e iteración
SCRUM	diseño iterativo, participación del usuario, pruebas continuas y creación de prototipos

Al tratarse del diseño de una interfaz avanzada es importante interactuar con el o los usuarios claves en el proceso de diseño. Para ello es importante trabajar con prototipos de diseño conceptual de lo que será la interfaz y el tipo de información que formará parte de la misma. Es importante considerar que es necesario realizar un diseño orientado a la CS que sea eje central en el proceso de DCU. Cabe destacar que, sin un análisis sistemático de los requisitos operacionales de los usuarios, sin una guía de factores humanos sólidos, y un programa de pruebas sistemático y objetivo, es casi imposible crear un sistema centrado en el usuario.

En [72] detalla que la CS solo existe en la mente del operador humano, presentar grandes cantidades de datos no servirá de nada a menos que la persona lo transmita, absorba y asimile de manera oportuna para formar la CS. Uno de los beneficios clave de observar la CS es que ésta le transmite al diseñador cómo deben combinarse todos los datos disponibles para facilitar la comprensión del usuario. Proporcionar sistemas que respalden el proceso de la CS es esencial para enfrentar un desafío crítico en los entornos complejos de la actualidad. En entornos complejos en donde la CS es un factor clave para la toma de decisiones determina que el desarrollo de software se centre en un enfoque estructurado que incorpore las consideraciones de la CS en el proceso de diseño, incluya la determinación de los requisitos de CS, el diseño para la mejora de CS y la medición y evaluación de la CS en el proceso de construcción.

7.3 Arquitectura del Framework. Middleware

El principal trabajo es diseñar la arquitectura de la capa intermedia, es decir, el *middleware*. El *middleware* está formado concretamente por las capas de *Módulos* y *Librerías*. Cada capa representa un nivel de abstracción en la arquitectura y está formada por un conjunto de componentes específicos. En su conjunto, entonces, las capas que forman parte de la arquitectura del framework están formadas por: las *aplicaciones* (aplicaciones que se desarrollarán tomando como base el framework) en la capa de más alto nivel, los *módulos* (es decir, formado por los componentes/módulos funcionales), *librerías* (bibliotecas de procesamiento de imágenes o visión artificial, manejo de mapas, comunicaciones, desarrollo de librerías propias, etc.) y el *sistema operativo*, en el nivel más bajo. Cada una de estas capas utiliza servicios ofrecidos por las anteriores, y ofrece a su vez los suyos propios a las capas de niveles superiores. De la misma manera, los componentes que forman parte del *middleware* se comunican entre sí ofreciéndose servicios entre ellos.

7.3.1 Capa: Módulos

Tomando como base el Modelo de Referencia para un Sistema de RA detallado en la sección 3.3 - *Partes de una aplicación de RA*, se diseñaron los componentes de software o módulos que forman parte del *middleware*:

- *Input*: Este módulo permite que se puedan ingresar datos al sistema, ya sea en forma de voz o gestos. La entrada vocal es útil a la hora de dar instrucciones simples al sistema (acciones que no requieran una comprensión total del entorno), por ejemplo, mostrar u ocultar un mapa, mostrar la ruta recorrida, establecer algún recordatorio o punto de referencia, etc. Este módulo cuenta con las clases relacionadas con el reconocimiento de los gestos y de la voz.
- *Location/Tracking*: Este módulo provee una interfaz unificada para la obtención de información de georeferenciación del dispositivo y al mismo tiempo devuelve la información (si se encuentra disponible) del sensor magnético y de la unidad de medición inercial (IMU). Permite realizar el tracking del usuario a través del uso de los sensores del dispositivo.

- *Mapping*: El módulo de mapeo implementa una interfaz a un SIG completamente funcional que permite cargar cartas prediseñadas a la aplicación (creadas con cualquier aplicación GIS compatible como ser ArcGIS, QGIS, OSM, GRASS, etc) y ubicar elementos fácilmente para su posterior visualización y manipulación ya sea en una carta 2D como en un mundo 3D.
- *Communications*: El módulo de comunicaciones consta de una simple abstracción que permite a dos o más dispositivos comunicarse entre sí mediante mensajes estructurados. El objetivo es testear la comunicación entre un cliente (implementado en JAVA-Android) y un servidor (implementado en JavaScript-Node.js). Como tecnología de comunicación entre estos componentes se utiliza Socket.io
- *Visualization*: Representación gráfica del lado del Cliente (dispositivo móvil) de la simbología u objeto renderizado. Radar 360 grados con alertas y amenazas, mapas de operaciones, estados del dispositivo, etc., son representados visualmente en la pantalla del dispositivo móvil. El filtrado de información prioritaria es procesado en este módulo para su visualización.
- *Vision*: El objetivo de este módulo es brindar las implementaciones de las clases necesarias para el reconocimiento de objetos. Se implementarán diferentes algoritmos, por ejemplo, detección y extracción de características, de matching (emparejamiento) y de clasificación. La finalidad es brindar al desarrollar de aplicaciones la mayor cantidad de algoritmos para que pueda implementarlo en el desarrollo de las aplicaciones correspondientes.

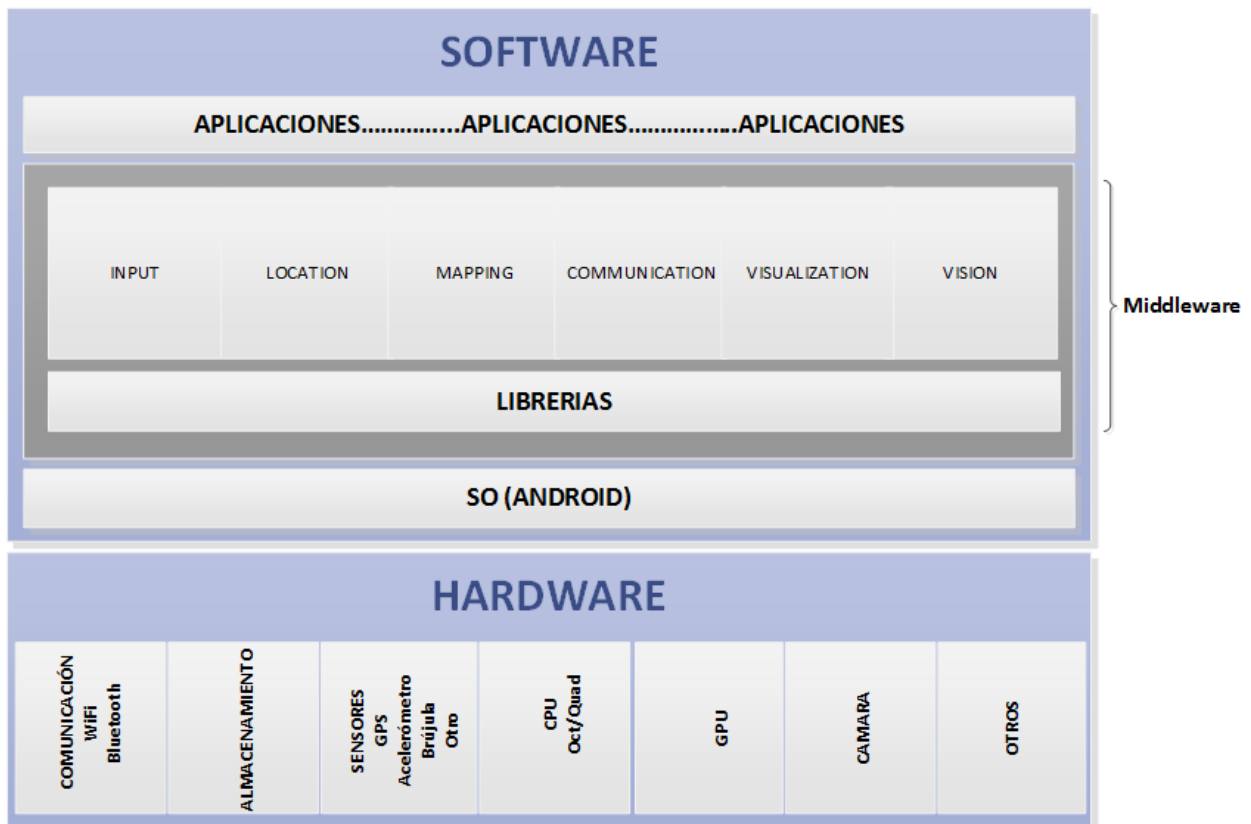


Figura 7.2 – Diseño basado en componentes de la arquitectura del framework (fuente: propia)

Por otro lado, en la siguiente tabla, se reflejan las contribuciones al logro de la consciencia de cada una de las capas que conforman la estructura del framework, es decir, se relaciona la caracterización de la CS y los módulos propuestos en el framework tomando como referencia las preguntas sugeridas por Gutwin et al. (ver sección 6.3 - *Identificación, selección y clasificación de requerimientos de la CS*).

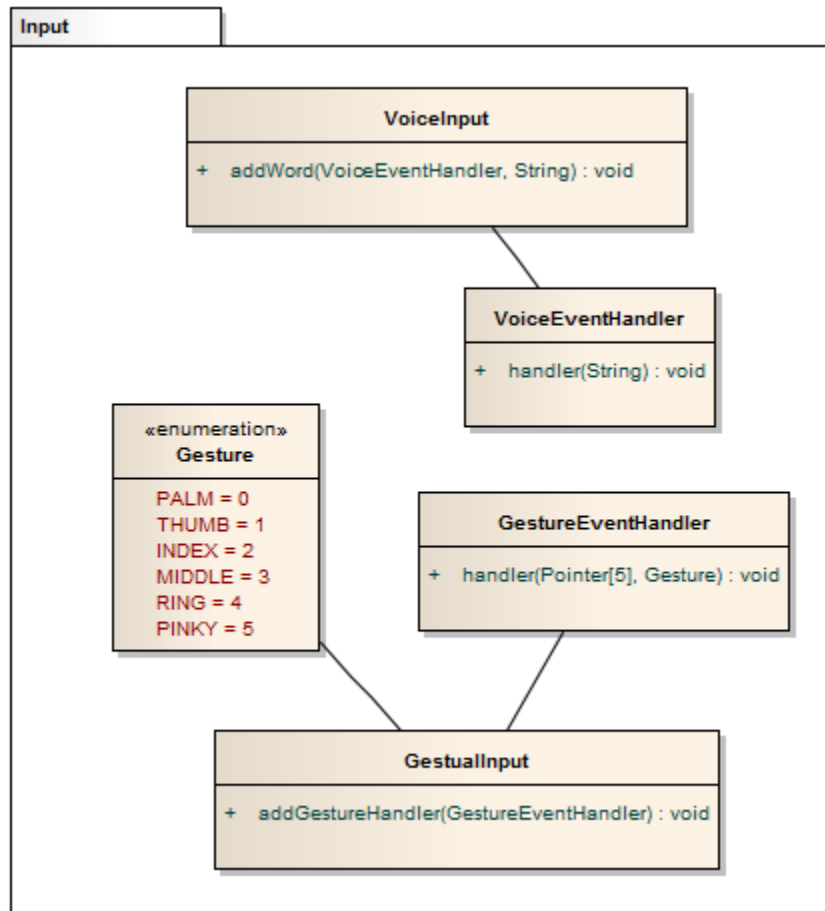
Tabla 7.2 Relación explícita entre la caracterización de la CS y los módulos propuestos

Módulos propuestos	Caracterización de la CS		
	<i>Quién</i>	<i>Qué</i>	<i>Dónde</i>
Input		X	
Location			X
Mapping			X
Communication		X	
Visualization	X	X	X
Vision	X	X	

A continuación, se detallan las características principales de cada módulo implementado en la Capa Módulos:

7.3.1.1 Módulo Input

Para que un sistema de RA sea útil en la vida real se necesita que interactúe con su entorno. Para ello se implementa el módulo de entrada de datos (módulo *Input*). Como se mencionó anteriormente este módulo permite ingresar datos al sistema, ya sea en forma de voz o gestos sin necesidad de utilizar un periférico externo. La entrada vocal es útil a la hora de dar instrucciones simples al sistema (acciones que no requieran una comprensión total del entorno), por ejemplo, mostrar u ocultar un mapa, mostrar la ruta recorrida, establecer algún recordatorio o punto de referencia, etc. Para instrucciones más complejas, como por ejemplo identificar un objeto específico, manipular un objeto virtual, etc; es necesaria una entrada mediante el uso de gestos, por lo tanto, para estos casos se provee la alternativa de entrada gestual. El framework detecta diferentes gestos manuales y los mapea a acciones que pueda manejar la aplicación.



**Figura 7.3 – Diagrama de Clases del módulo Input
(fuente: propia)**

7.3.1.2 Módulo Location

En un sistema de RA integrado a GIS suele ser importante poder obtener datos de geolocalización del usuario. Este módulo está basado en la obtención de información del GPS, del sensor magnético y de la unidad de medición inercial que es un dispositivo electrónico que mide e informa acerca de la velocidad, orientación y fuerzas gravitacionales del dispositivo, usando una combinación de acelerómetros y giróscopos. Estos sensores están presentes en la mayoría de los dispositivos móviles inclusive en las gafas de RA.

De esta forma una aplicación que utilice el framework tiene acceso no sólo a la ubicación geográfica del usuario sino también su posición y orientación en 3 dimensiones.

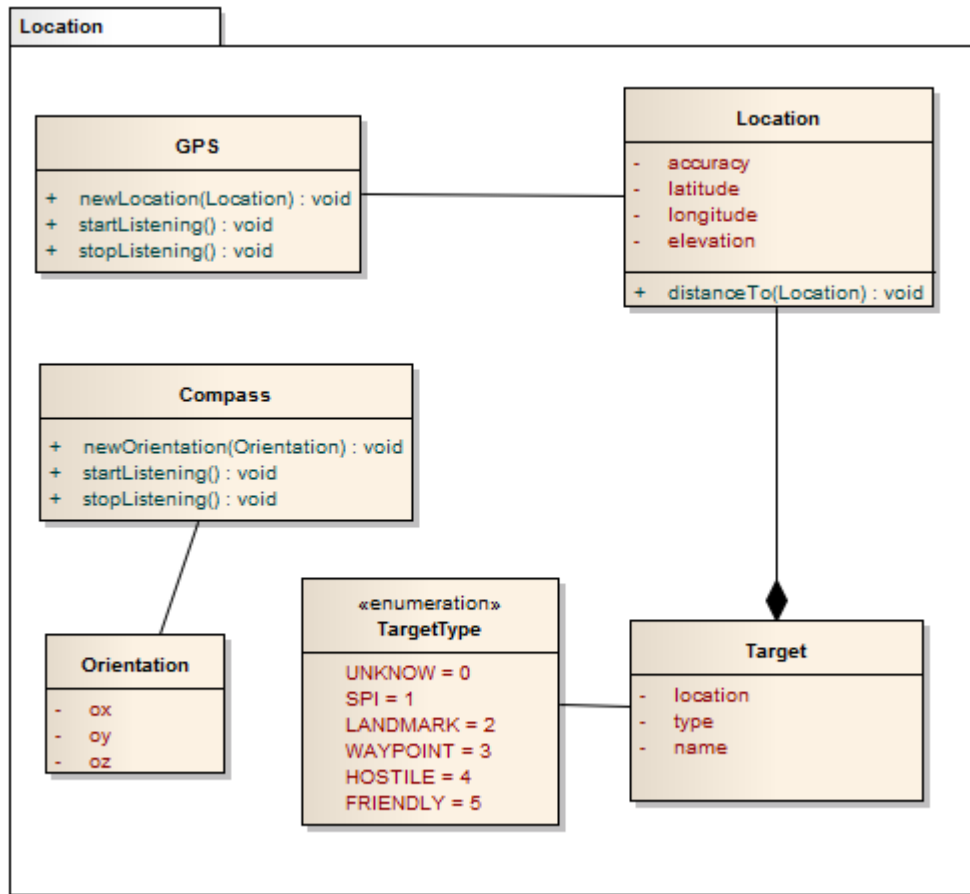


Figura 7.4 – Diagrama de Clases del módulo Location (fuente: propia)

7.3.1.3 Módulo Mapping

Los GIS posibilitan la ubicación espacial de objetos e una carta. Estos sistemas son fundamentales para el correcto desarrollo de actividades militares, por ejemplo, de inteligencia. Si bien la acción de “ubicar elementos en una carta” parece trivial, existen muchas dificultades a la hora de hacerlo correctamente.

El sistema en su estado actual permite la carga de elementos georreferenciados (ya sea cargados de antemano en el código o dinámicamente en tiempo de ejecución) y cartas 2D en formato OSM (OpenStreetMaps). Si bien ninguno de los programas GIS de uso general

son capaces de exportar cartas en formato OSM, la aplicación *Osmosis*⁹ puede ser utilizada para convertir casi cualquier formato de los más conocidos (por ejemplo shp) al formato necesario. Se pueden utilizar tanto cartas *raster* como *vectoriales*.

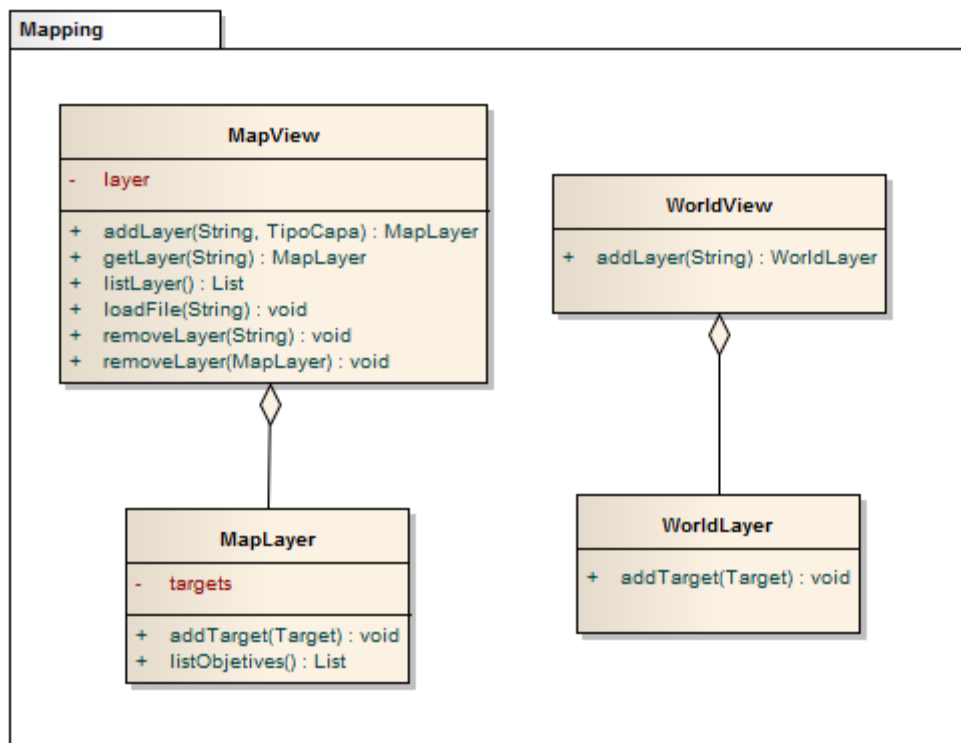


Figura 7.5 – Diagrama de Clases del módulo Mapping
(fuente: propia)

7.3.1.4 Módulo Communication

Las comunicaciones entre los diferentes dispositivos son muy importantes en operaciones militares. Es necesario estar comunicados con el JPat ya que permite que el resto del comando siga sus instrucciones y órdenes en tiempo real. Por otro también es vital la información que es enviada desde el CII y desde este al operador en avanzada. Una demostración del uso de las comunicaciones entre usuarios utilizando la librería *Libstreaming* detallado anteriormente. En la prueba de concepto inicial que se describe en la sección 8.4 - *Módulo*

⁹ Osmosis: <http://wiki.openstreetmap.org/wiki/Osmosis>

Communication se realizó la prueba de uso de este módulo implementado en una aplicación móvil cliente / servidor.

El sistema de comunicaciones funciona perfectamente para comunicaciones *punto a punto* e incluso para conexiones *Multicast*.

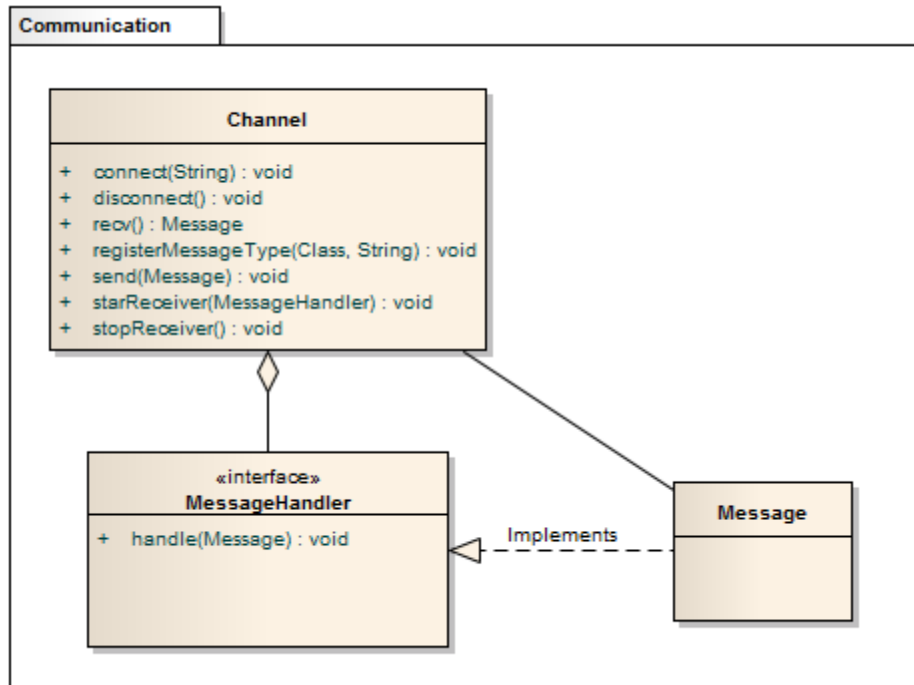


Figura 7.6 – Diagrama de Clases del módulo Communication (fuente: propia)

7.3.1.5 Módulo Visualization

El filtrado de información prioritaria corresponde a aquellos datos que el operador puede visualizar como primera capa de información contextual. El objetivo principal es de no abrumar al operador con abundantes datos del contexto ya que lo puede llevar a perder la concentración en las tareas que está ejecutando. De acuerdo a lo expresado en la sección 6.3 - *Identificación, selección y clasificación de requerimientos para la CS*, y complementando con lo detallado por [73], el DCU no sólo significa presentar a los usuarios la información que necesitan en un momento dado sino que un intento bien intencionado para ayudar a administrar el problema de sobrecarga de información es la idea del filtrado de información.

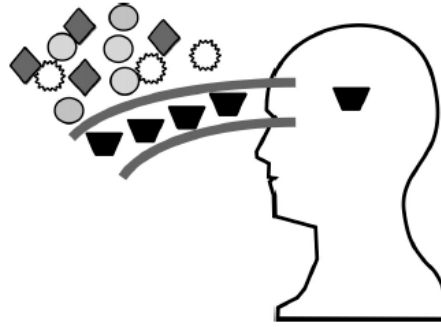


Figura 7.7 – Filtrado de información. Cierta conjunto de información es prioritario. Se excluye el resto de información (fuente: Endsley at al. [73])

Tomando los requisitos de usuarios mencionados en la sección 6.3 - *Identificación, selección y clasificación de requerimientos de la CS*, en el framework propuesto se implementó la funcionalidad para que soporte hasta 2 capas de información. El diseño y disposición de los objetos en la pantalla se desarrolló siguiendo el criterio de DCU focalizado en la CS de los usuarios tal cual se detalló en la sección mencionada con anterioridad.

La primera capa de información se visualiza siempre directamente en la pantalla del dispositivo móvil Cliente.

En cuanto a la segunda capa, la información es solicitada a demanda por el operador ya sea a través de reconocimiento gestual o vocal.

Filtro de información 1:

La primera capa de información está compuesta por:

- Radar 360 grados con información de alertas y amenazas
- Ubicación del operador y de los soldados en el Radar
- Información de geolocalización
- Comunicación con el CII
- Información del estado general del dispositivo e información del lugar y fecha

En la siguiente figura se detalla la estructura de visualización del radar de 360 que consta de una elipse que representa la orientación del usuario. El observador u operador se ubica en el centro de dicha elipse. Un punto rojo marca el norte y sobre la línea verde se sitúan los objetivos precargados en el sistema (o recibidos mediante mensajería). Por otro lado el radar

está provisto de información de geolocalización y, además, posee los límites del campo visual en donde los objetivos son mostrados en el campo visual del observador. Los elementos dentro del campo visual son representados espacialmente en pantalla mientras que el resto son ocultados.

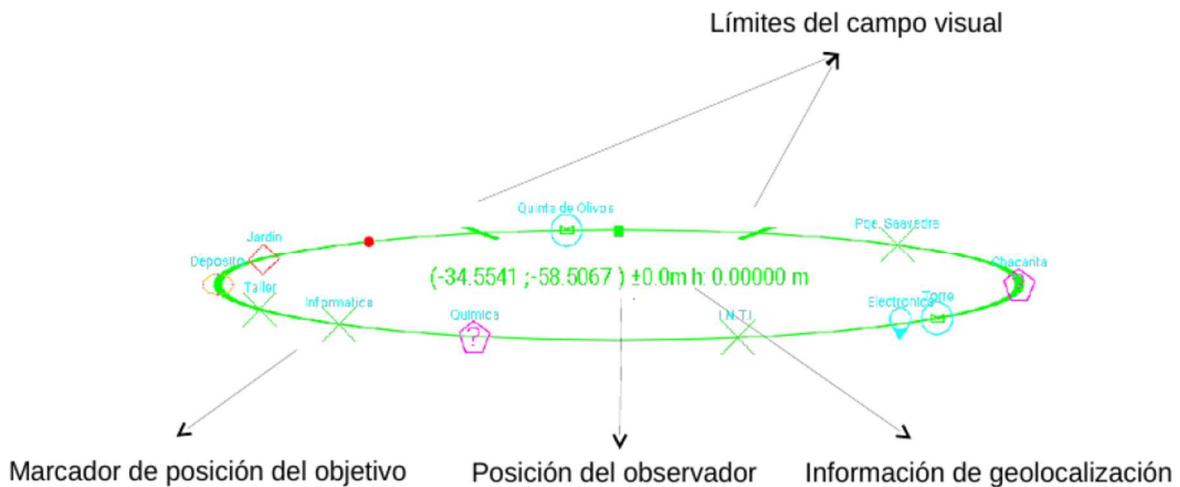


Figura 7.8 – Simbología del radar. Estructura de la visualización del radar (fuente: propia)

El radar es la parte central del sistema y es lo que facilita la navegación, marcando la dirección y distancia de los distintos elementos que el sistema reconoce. Utilizando la API de localización, provista por el módulo *Location*, el sistema obtiene las coordenadas GPS del usuario y al mismo tiempo su orientación y velocidad de desplazamiento. Esta información es la que permite calcular los diferentes rumbos y distancias necesarios para construir la visualización.

En la siguiente figura se puede apreciar que en la parte inferior de la pantalla se encuentra el radar de 360 grados con el uso de simbología (iconografía) específica.

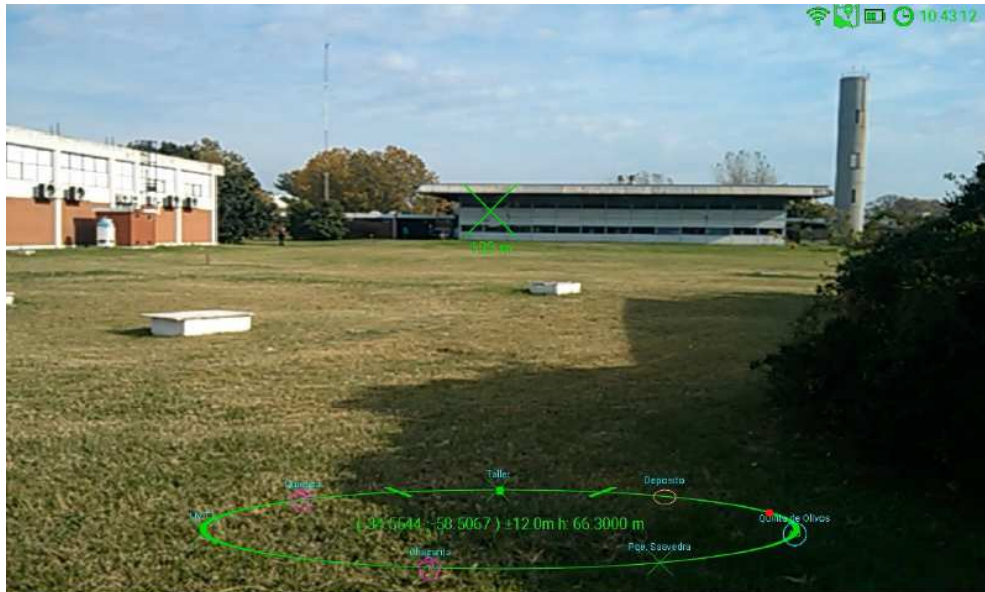


Figura 7.9 – Mapa 360 grados del Prototipo desarrollado utilizando el framework (fuente: propia)

Uno de los principales problemas o limitaciones de este radar es la precisión que puede obtener un GPS de dispositivo móvil. La precisión de la ubicación siempre es mayor a los 10 metros, haciendo que la ubicación de los objetivos cercanos sea poco confiable y experimenten cambios repentinos durante la marcha. Para mitigar estos errores se utiliza un *filtro pasabajos* en la ubicación del GPS, evitando así los saltos bruscos. No es una solución al problema real, pero evita los saltos repentinos de los objetos, brindando una experiencia más amigable al usuario.

Dentro de la simbología militar utilizada se dispuso un conjunto de símbolos con el objetivo de probar el fácil intercambio de éstos en la configuración de la aplicación. Para una mejor interpretación de los mismos se los diferencié del tipo de alerta / amenaza.

Filtro de información 2:

La segunda capa de información está compuesta por:

- Mapas del área de operaciones
- Información contextual ampliada
- Reconocimiento de objetos
- Radar de alertas y amenazas

La capa de información siguiente está compuesta por acciones que se solicitan a demanda por el usuario a través de comandos gestuales o vocales. El framework soporta hasta 5 comandos gestuales diferentes cada una asociada a un dedo de la mano.



Figura 7.10 – Simbología gestual
(fuente: propia)

En la siguiente figura se detalla el diagrama de clases del módulo *Visualization*:

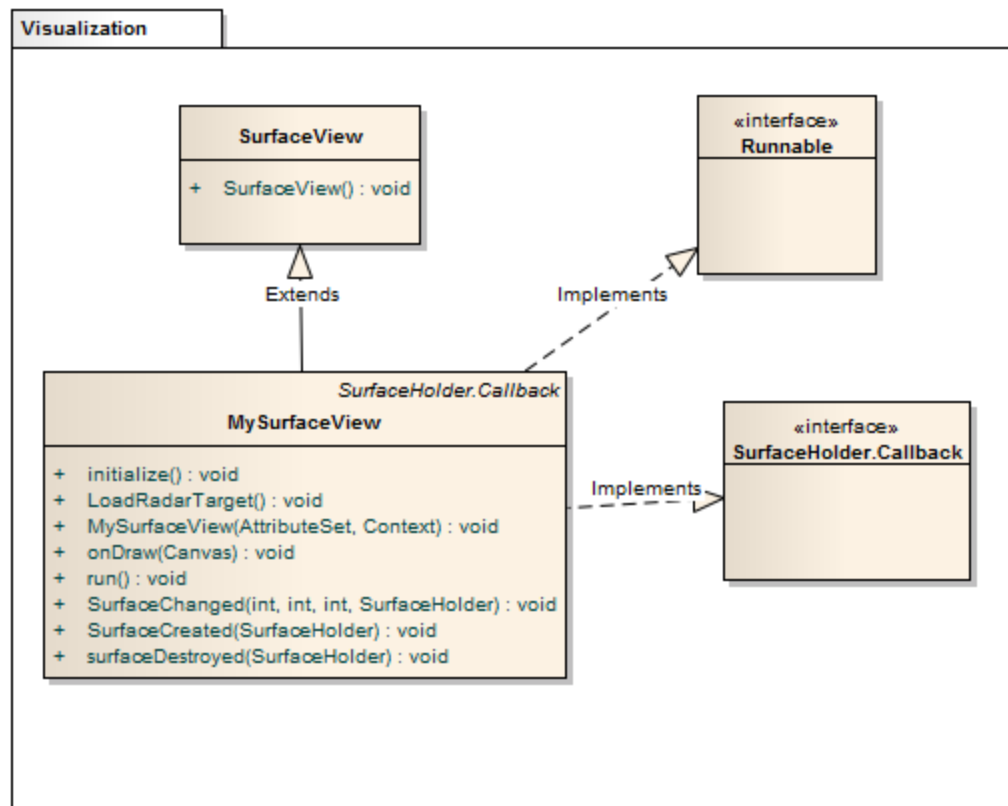


Figura 7.11 – Diagrama de Clases del módulo Visualization
(fuente: propia)

En la siguiente figura se presenta el diseño completo de la estructura de paquetes / módulos del framework propuesto:

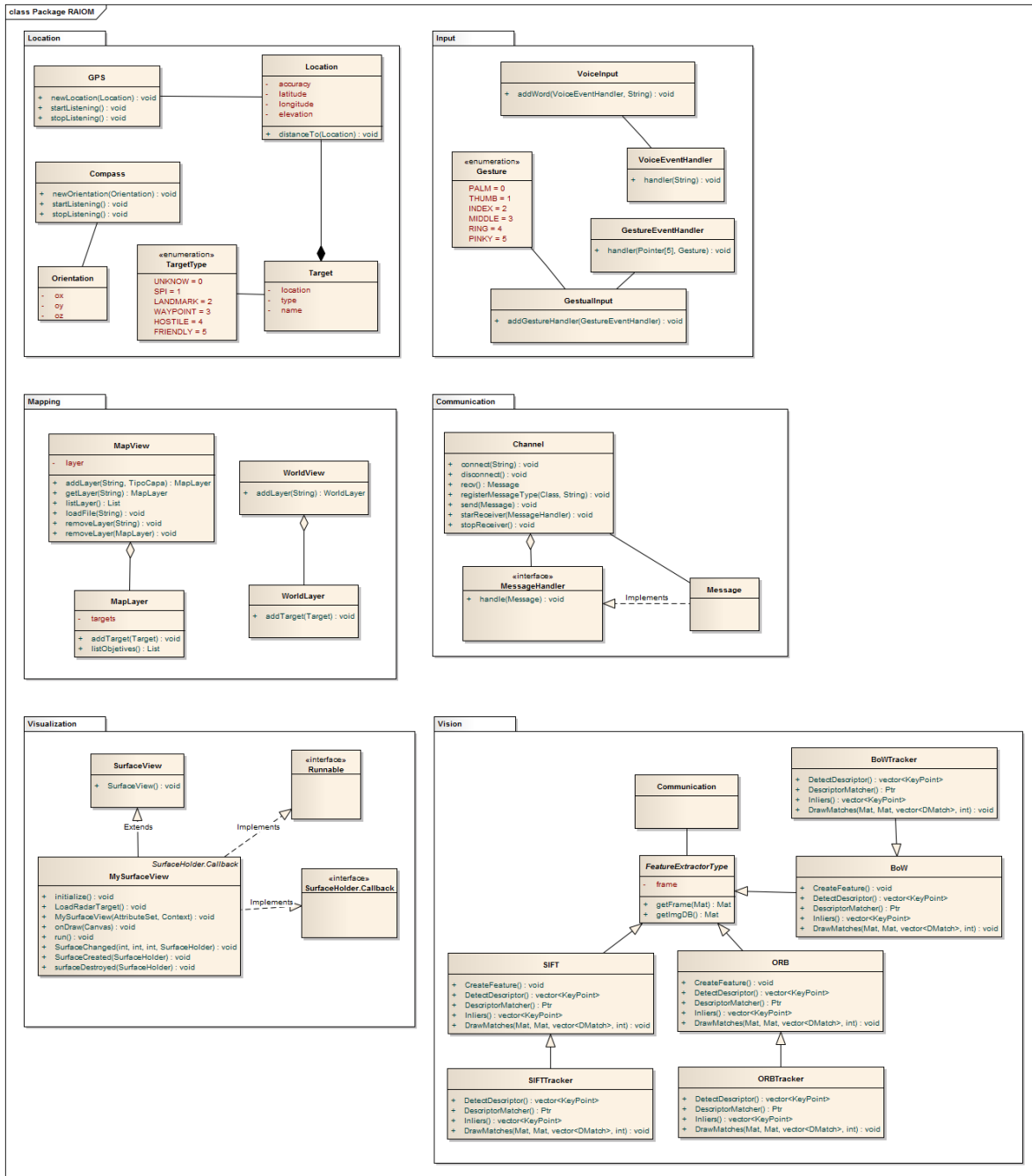


Figura 7.12 – Diagrama de Clases de los módulos del framework (fuente: propia)

Una de las características más relevantes de la arquitectura del framework es que el diseño arquitectónico soporte el despliegue de las aplicaciones que se desarrollen (y utilicen

al framework) en un ambiente distribuido del tipo Cliente / Servidor. Más detalles en la sección 7.5 - *Proceso distribuido*. Se decidió que el framework sea desarrollado bajo el sistema operativo Android pues al ser de código abierto permitirá tener mayor libertad de integración de las clases de este marco de trabajo con los recursos de las capas más bajas. En cuanto al lenguaje de programación principal utilizado en el diseño del framework será Java ya que estará basado en Android, en el lado del Cliente y Python en el lado del Servidor. El framework utilizará OpenCV como librería de procesamiento de imágenes tal cual se mencionó en la sección 5.6.2 - *Optimización a nivel software*.

A continuación, se detalla el módulo de *Vision* que es el más complejo y más extenso del framework propuesto:

7.3.1.6 Módulo Vision

El módulo de *Vision* y el módulo de *Input* se apoyan en el uso de la librería OpenCV que se encuentra en la capa de *Librerías* (ver sección 7.3.2 - *Capa: Librerías*) del framework propuesto.

Como se menciona en la sección 7.5 - *Proceso distribuido*, el procesamiento de imágenes se realiza del lado del Servidor ejecutándose sobre un dispositivo ODROID-XU3. Se han experimentado con diversos algoritmos de extracción de características con el objetivo de reconocer objetos determinados y se ha evaluado tanto su nivel de invariancia como de performance. Se tomó como referencia para determinar cuál o cuáles algoritmos son los más adecuados las características del ambiente. Estas características estaban determinadas por la luminosidad del entorno, tamaño y rotación de los objetos.

El framework posee un componente de entrada (módulo *Input*) que le permite al usuario ingresar datos a través del reconocimiento gestual y vocal. Con gestos simples se le puede indicar al dispositivo que se quiere ejecutar uno de varios comandos posibles. Al colocar la palma de la mano abierta frente a la cámara se presentan hasta un máximo de 5 opciones de menús. Para elegir una opción basta con flexionar el dedo en cuestión, pasando a otro menú o ejecutando una acción específica. Existen varios métodos de reconocimiento gestual. Todos presentan un paso de segmentación, y la mayor parte de ellos utiliza segmentación por color. La segmentación por color consiste en buscar en la imagen todas las regiones que contienen colores parecidos a la piel. Luego de encontradas las regiones se

analizan con técnicas que varían desde reconocimiento por forma hasta reconocimiento por clasificador de imágenes utilizando características o PCA (Principal Component Analysis). El problema central de casi todos los algoritmos en esta área es que están pensados o desarrollados para ser utilizados en ambientes controlados. Cuando se utilizan en ambientes exteriores y/o con movimientos bruscos de cámara los resultados son pobres. Para el reconocimiento gestual se utiliza una técnica de segmentación adaptativa por color. Esta técnica requiere una calibración inicial con una imagen de la mano en cuestión. La calibración extrae un clúster de colores de la mano. El procedimiento incluye una conversión de la imagen al espacio de color HSV. Mediante esta conversión se obtienen los rangos de HS (Hue y Saturation) que nos interesan para la mano, admitiendo cambios en V (Value) que permite que el reconocimiento sea robusto ante cambios de brillo.

Para realizar la segmentación de la mano se aplica la misma transformación a HSV y se buscan los sectores de la imagen con el rango HS adecuado. Se extraen los sectores candidatos (blobs) y se analiza la forma del contorno obtenido. Si se puede considerar que el sector corresponde a una mano se sigue con el procedimiento utilizando un algoritmo de *convex hull* y *convexity defects* que sirve para detectar la región de los dedos.

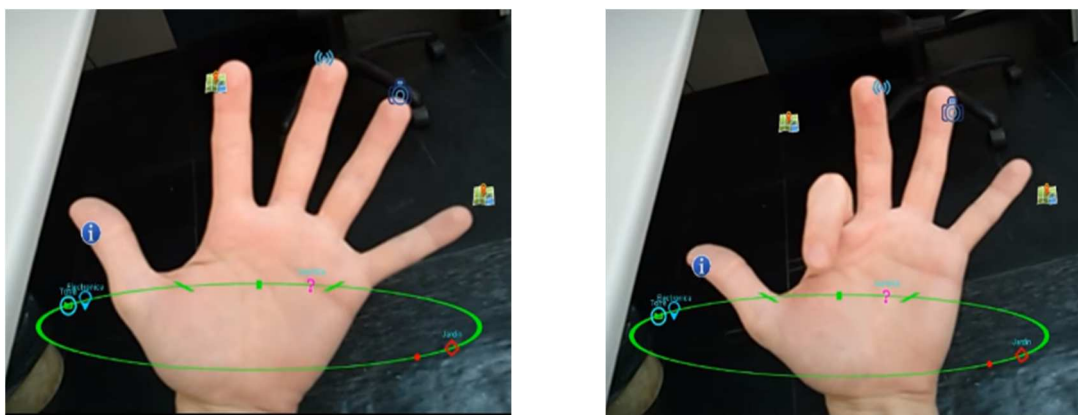


Figura 7.13 – Pruebas iniciales de reconocimiento gestual y visualización de iconos (fuente: propia)

El reconocimiento de objetos se lleva a cabo mediante técnicas de emparejamiento de características. Estas técnicas consisten en extraer un conjunto de características de una serie de imágenes de entrenamiento del objeto en cuestión. Cada conjunto de características está asociada al objeto de cuya imagen fueron extraídas.

Una vez extraídas todas las características se arma un modelo Bag of Word (BoW) [74], que es una técnica de clasificación de documentos aplicado a clasificación de imágenes. Cada imagen es tratada como un documento, mientras que cada característica como una palabra. El conjunto de todas las características compone el vocabulario. Con este vocabulario se arma un índice invertido o un hash que permita la recuperación de las imágenes a las que pertenece. BoW es un algoritmo utilizado originalmente para clasificación de documentos de texto. Cada documento es considerado una bolsa de palabras, donde no importa el orden de cada una. En procesamiento de imágenes se utiliza de manera similar, considerando a cada grupo de imágenes que represente un objeto como un documento, y por lo tanto como una bolsa. En vez de utilizar palabras se utilizan las características extraídas de la imagen mediante algún método de extracción de características como SURF, SIFT, FREAK o FAST.

Para realizar el reconocimiento se extraen las características de la imagen en cuestión y se realiza la búsqueda en el índice. La técnica entonces permite encontrar imágenes similares y además clasificar imágenes según su parecido con otras.

En la siguiente figura se ilustra el modelo utilizado. Las características (features o landmarks) son descripciones de vértices, esquinas o algún otro punto saliente de una imagen. En definitiva, son una descripción de puntos particulares (puntos clave, keypoints o interest points) que pueden ser fácilmente distinguidos (estas descripciones en la práctica son vectores multidimensionales).

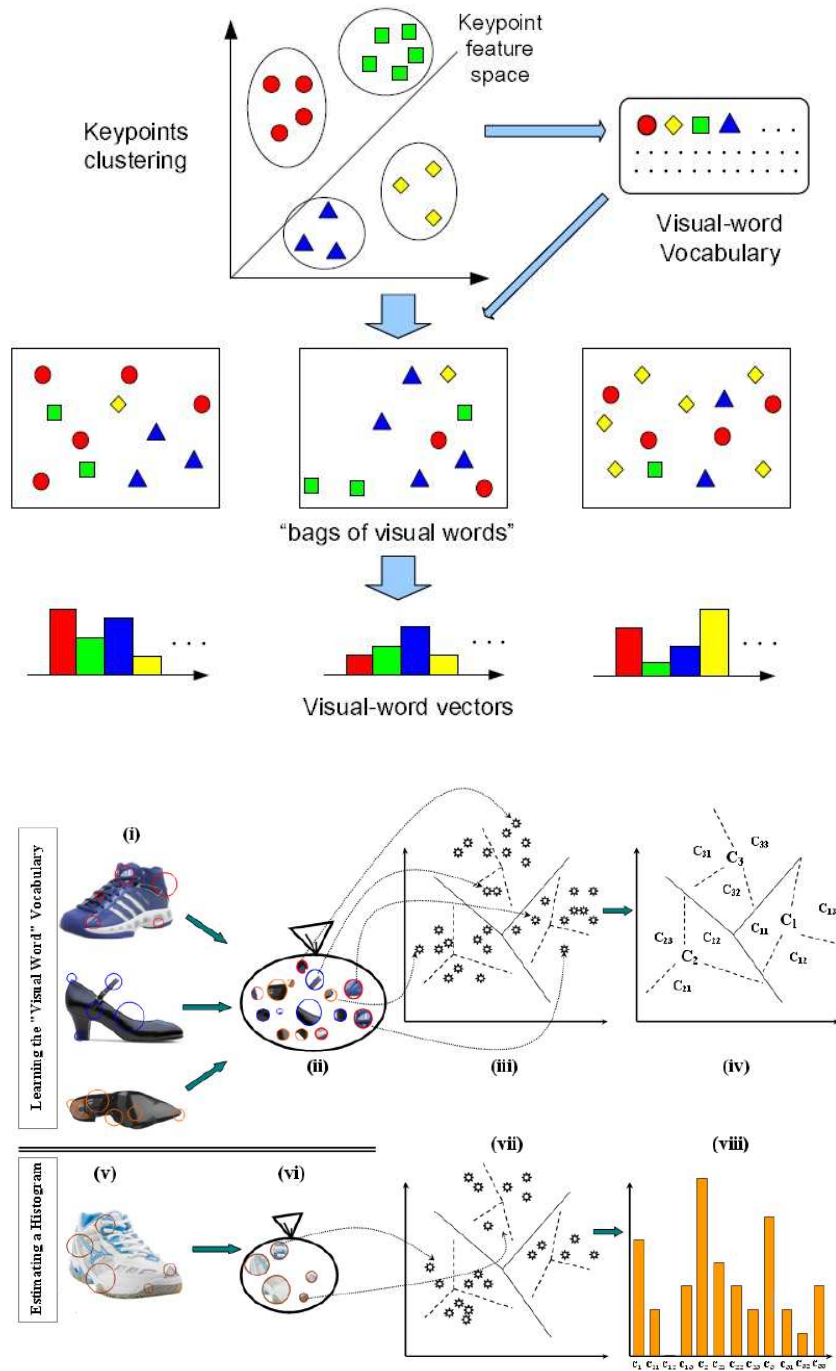


Figura 7.14 – Explicación gráfica de Bag Of Words: (i) imágenes representando una palabra visual (ii) características (iii) descriptores ubicados en el espacio (iv) agrupación en clusters (v) imagen a clasificar (vi) características (vii) correspondencias con los descriptores conocidos (viii) histograma de correspondencias (fuente: Jun Yang et al [74])

En general se quiere que estas características sean reconocibles aun habiendo sufrido la imagen cambios en el punto de vista, rotaciones o cambios de iluminación (es decir, que sean invariantes ante tales cambios).

Existen muchos algoritmos diferentes tanto para búsqueda de puntos clave como para extracción de características.

Del lado de la búsqueda de puntos clave se encuentran algoritmos como los detectores de esquinas de Harris y Shi-Tomasi, los detectores LoG (Laplacian of Gaussian), DoG (Difference of Gaussians), DoH (Determinant of Hessian), FAST (Features from Accelerated Segment Test), PCBR (Principal Curvature-Based Region detector) o MSER (Maximally Stable Extremal Regions) entre otros. Algunos de los algoritmos son más estables que otros ante cambios de brillo y orientación, algunos son más lentos que otros, pero en definitiva ninguno es perfecto y el uso de cada uno depende en gran parte de una elección personal y el compromiso que se pueda tomar entre velocidad y precisión necesaria [75].

Por el lado de los descriptores de puntos clave se encuentran algoritmos como SIFT, SURF, ORB, GLOH y LESH entre otros. Cada uno de los algoritmos realiza una o varias operaciones en los puntos clave encontrados.

SIFT [76] por ejemplo (el más popular y preciso, pero a la vez el que más proceso computacional requiere) calcula histogramas orientados en vecindades de 4x4 píxeles dentro de un parche de 16x16 píxeles en torno al punto clave. En total calcula 16 histogramas con 8 posibles orientaciones cada uno, resultando en un vector de características de 128 elementos.

Por otro lado, SURF utiliza sumas de transformadas Wavelet Harr, resultando en un algoritmo mucho más rápido. En este trabajo se optó por utilizar el detector FAST con el descriptor FREAK.

FAST [77] es ideal para aplicaciones de RA en dispositivos móviles gracias al escaso tiempo de procesamiento requerido. Es significativamente más rápido que los detectores tradicionales como DoG o Harris.

FREAK [78] es un descriptor que intenta imitar el funcionamiento del ojo humano. Utiliza un patrón de muestreo centrado en el punto clave que consiste en zonas circulares con un incremento en el tamaño y superposición de las zonas a medida que se alejan del centro, similar a la distribución de células receptoras en el ojo: alta concentración de conos (muy sensibles) en el centro y luego en disminución hasta quedar solo bastones (menos sensibles).

Esta forma de muestreo permite tener mejor resolución en el centro (el punto de interés) pero sin perder los detalles que lo rodean.

Adicionalmente el algoritmo introduce una búsqueda rápida de correspondencia: búsqueda sacádica. Esta técnica también intentando imitar la forma en la que funciona el ojo: el ojo no se queda generalmente fijo en un punto, si no que recorre la escena con movimientos rápidos (sacádicos).

El algoritmo busca correspondencias primero comparando la parte del descriptor correspondiente a la zona de menor resolución, si esta zona presenta cierta similitud dentro de márgenes razonables el algoritmo continúa la comparación con los niveles de mayor resolución. Si la zona de menor resolución no presenta parecido, pasa a otro descriptor. De esta forma los descriptores sin correspondencias se descartan rápidamente.

En contrapartida a todas las ventajas que presenta, FREAK tiene la desventaja de no ser tan robusto ante cambios de escala para factores de 2.5 en adelante como lo son otros descriptores como SIFT, SURF o BRIEF.

En la siguiente figura se muestra una prueba de concepto implementada en un dispositivo móvil para la identificación de objetos utilizando el algoritmo ORB.

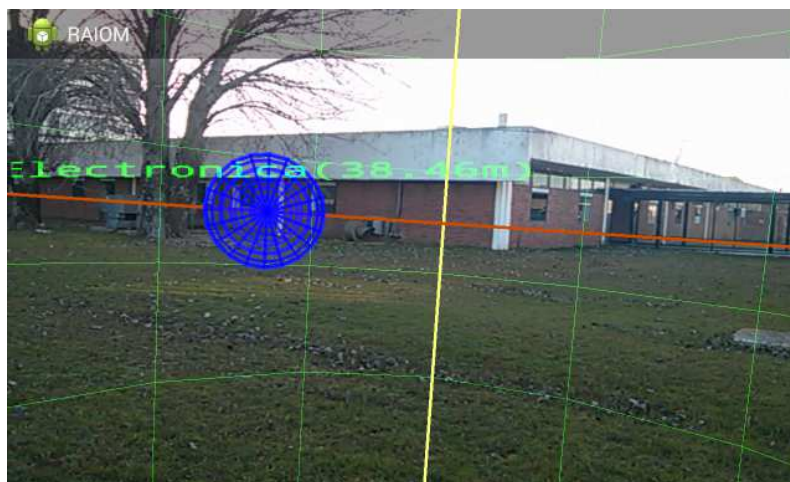


Figura 7.15 –Prueba de identificación del edificio de Electrónica del CITEDEF utilizando ORB (fuente: propia)

A continuación se detalla el diagrama de clases correspondiente al módulo *Vision*:

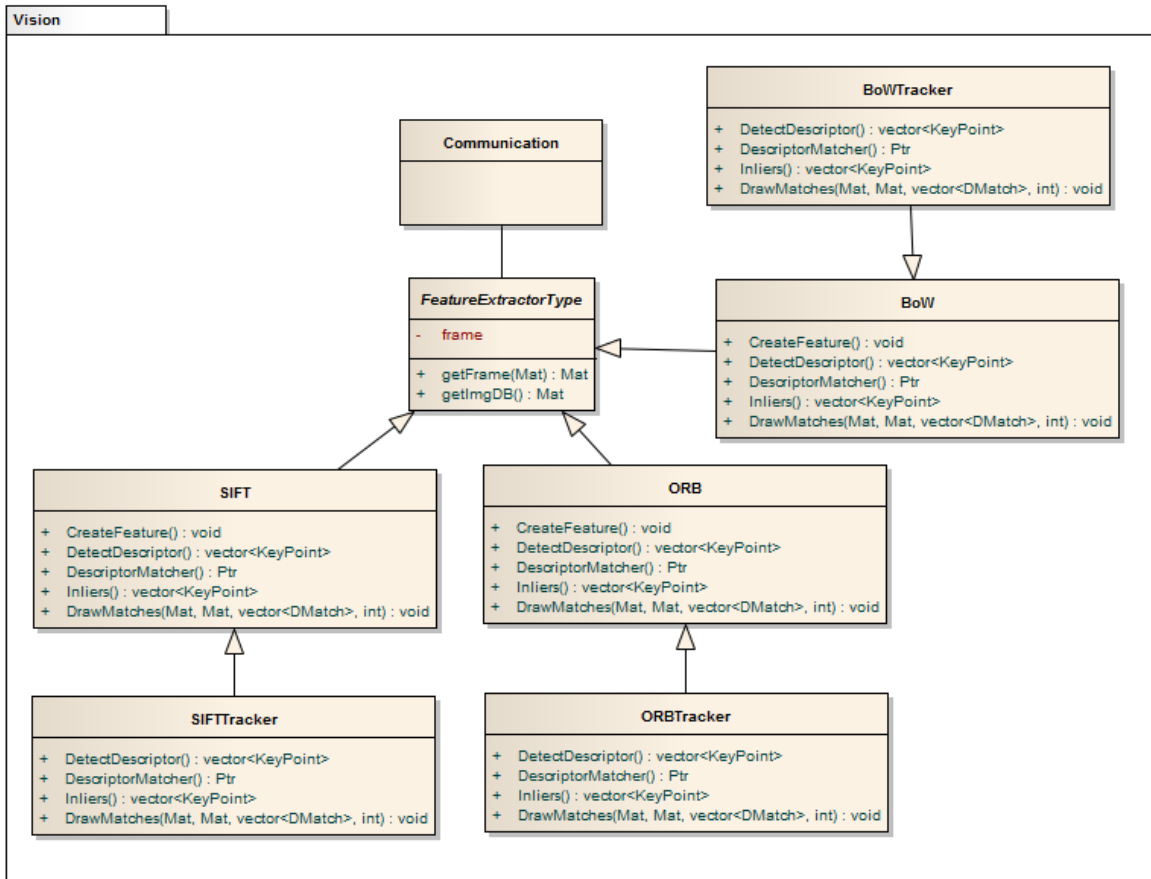


Figura 7.16 – Diagrama de Clases del módulo Vision (fuente: propia)

7.3.2 Capa: Librerías

Por otro lado, en la capa de *Librerías* se han integrado al framework propuesto un conjunto de librerías entre las que se destacan:

- *OpenCV*: Explicado en la sección 5.6.2 - *Optimización a nivel software*
- *Mapsforge*: ofrece una librería libre, offline y de código abierto para el sistema operativo Android. El objetivo es la creación de aplicaciones basadas en OpenStreetMap¹⁰ (OpenStreetMap es un proyecto perteneciente a la comunidad open source que permite crear y editar mapas de manera libre). Las herramientas y APIs

¹⁰ OpenStreetMap: <https://www.openstreetmap.org/>

proporcionadas incluyen soluciones para representación en mapas, planificación y navegación de rutas, indexado y búsqueda de POI, capas en los mapas entre otras opciones. También soportan capas (POIs y polígonos arbitrarios).

- *Libstreaming*: es una API que permite realizar *stream* de la cámara y/o del micrófono para aplicaciones que se ejecutan en el sistema operativo Android utilizando el protocolo RTP¹¹.
- *Socket.io*¹²: permite la comunicación bidireccional en tiempo real basada en eventos. Funciona en todas las plataformas, navegadores o dispositivos, centrándose igualmente en la fiabilidad y la velocidad.
- *ZeroMQ*¹³: esta biblioteca utiliza tecnología de comunicación sobre sockets TCP/IP. Además, es un sistema completo de colas de mensaje, concurrencia, balanceo de cargas, enrutamiento muy eficaz y rápido que abre la oportunidad para crear aplicaciones de cómputo distribuido. Es una tecnología GPL y de código abierto.
- *PocketSphinx*¹⁴: es un motor ligero de reconocimiento de voz, específicamente adaptado para dispositivos de mano y dispositivos móviles.

A continuación, en el siguiente gráfico se ilustra la vista de componentes y conectores

¹¹ RTP: Real-time Transport Protocol. Es un protocolo de nivel de sesión utilizado para la transmisión de información en tiempo real) sobre UDP.

¹² Socket.io: <https://socket.io/>

¹³ ZeroMQ: <http://zeromq.org/>

¹⁴ PocketSphinx: <https://github.com/cmusphinx/pocketsphinx>

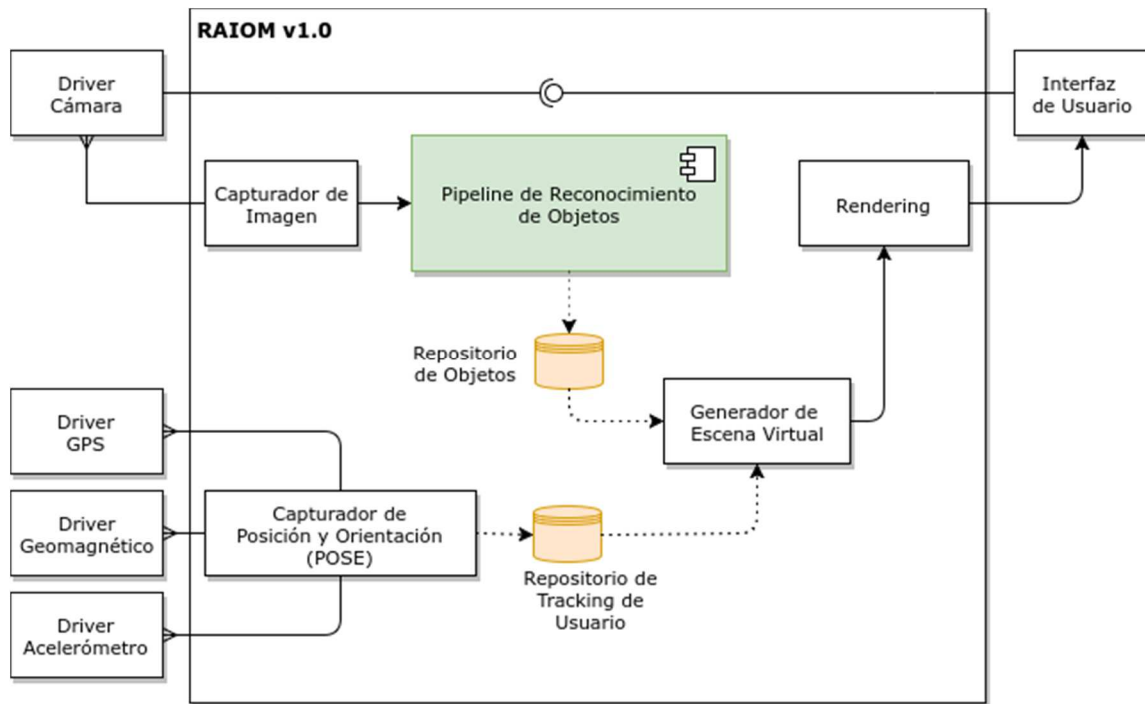


Figura 7.17 – Arquitectura propuesta del sistema utilizando la vista de componentes y conectores (fuente: propia)

7.4 Arquitectura del Sistema. Componentes de hardware utilizados

La arquitectura del sistema propuesta está representada por un conjunto de componentes hardware en dónde se ejecutarán las aplicaciones desarrolladas que extenderán el funcionamiento del framework propuesto en el presente trabajo de tesis. Estos componentes hardware son:

- *Smart Glasses Moverio BT-300 Smart Glasses:* gafas de RA donde integra sensores inerciales, GPS, brújula magnética, luminosidad. Estas gafas oficiarán de prototipo para la proyección y visualización de los elementos digitales desarrollados en las pruebas de concepto.
- *Cardboard box:* soporte para el dispositivo móvil que se encargará de proyectar y visualizar las imágenes sobreimpresas. Adicionalmente, de los dispositivos móviles se utilizarán los sensores que trae incorporado en la mayoría de los modelos, tales

como, sensores inerciales (acelerómetros y giróscopos), GPS, brújula magnética, entre los principales. Dichos dispositivos oficiarán de pantalla y, adicionalmente como parte de la arquitectura de componentes de hardware propuesta, se utilizarán los diferentes sensores IMU.

- *Procesador/es externo (ODROID XU3)*: este dispositivo se encargará de procesar las imágenes capturadas por la cámara. El resultado es enviado al dispositivo móvil (smartphone o smart glasses) para su proyección y visualización.
- *Cámara HD*: este componente capturará las escenas exteriores. Posteriormente las imágenes obtenidas serán procesadas por los dispositivos ODROID XU3.
- *Micrófono*: se utilizará para la entrada de comandos vocales mediante la utilización de reconocimiento de voz.
- *Sensores externos*: estos componentes son provistos por el Ejército y básicamente son todos aquellos dispositivos que toman datos del contexto (UAV, Cámaras, Magnéticos, etc.) y los datos son enviados al CII para su procesamiento.
- *Medio de comunicación*: en los prototipos desarrollados (aplicaciones que utilizan el framework diseñado) se utilizaron antenas Wireless para transmitir datos desde el CII y el JPat.

A continuación, en la siguiente figura se representa la Arquitectura del sistema sugerida en su totalidad. Se detallan los principales componentes de hardware propuesto:



Figura 7.18 – Arquitectura del Sistema de RA propuesto (fuente: propia)

El propósito del presente trabajo es centrarse en el diseño del framework para que las aplicaciones desarrolladas sean ejecutadas en dispositivos móviles con sistema operativo Android y en procesadores con arquitectura ARM. El framework deberá estar diseñado de manera tal que pueda paralelizar funciones de visión por computador en procesadores ARM distribuidos en la arquitectura de sistemas propuesto.

Cabe aclarar que en esta primera etapa se evaluó el uso de componentes de hardware del tipo COTS, ya que los mismos se encuentran muy probados por los usuarios finales y los costos de adquisición son muy inferiores a los componentes de hardware militares.

La arquitectura del sistema está pensada para que las aplicaciones que utilicen el framework propuesto en el presente trabajo puedan ejecutarse tanto en gafas de RA como también en dispositivos móviles tales como smartphone, tablets, etc.

Siguiendo la arquitectura planteada en la sección 4.3.3 - *Proyecto iARM* y en la arquitectura de sistemas detallado en la sección 4.3.4 - *Proyecto ULTRA-Vis*, se decidió desarrollar una arquitectura similar pero con procesamiento distribuido con el objetivo de optimizar las tareas de procesamiento de imágenes.

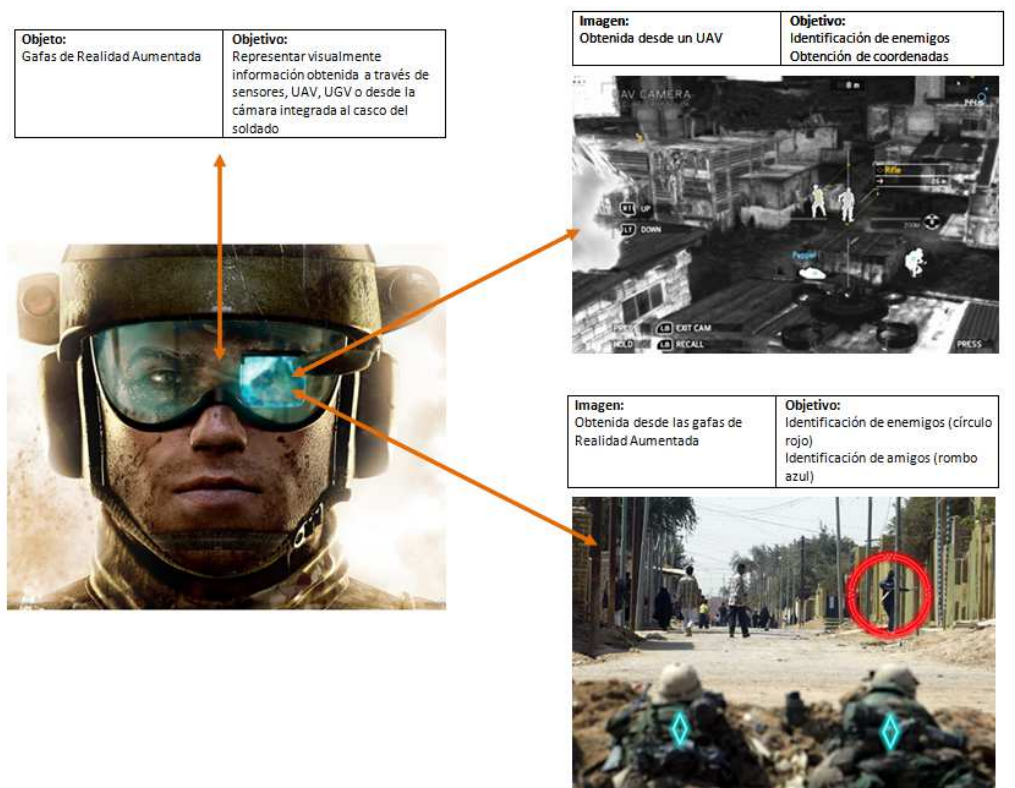


Figura 7.19 – Diseño de las capacidades del framework propuesto (fuente: propia)

Esta arquitectura responde a las necesidades de derivar carga de trabajo intensivo a procesadores externos bajo una arquitectura distribuida para, por ejemplo, procesar los algoritmos necesarios de visión por computador y devolver el resultado al dispositivo móvil que se encargará de proyectar el resultado correspondiente. La razón principal de enviar parte del procesamiento a los procesadores externos distribuyendo la carga se debe a que los componentes móviles cuentan con procesadores de bajo poder de cómputo con lo cual dificulta obtener resultados al procesar gran volumen de información en un tiempo razonable.



Figura 7.20 – Ejemplo de la Arquitectura propuesta utilizando gafas de RA (1) y ODROID XU3 (2) (fuente: propia)

Por lo tanto, para ejecutar el procesamiento distribuido se optó por los procesadores ODROID XU3 de la empresa *Hardkernel co., Ltd* de Corea del Sur. Estos componentes de hardware fueron los seleccionados para formar parte de la arquitectura integral del sistema y cuya función será la de procesar externamente imágenes y otras funciones. Se escogió este procesador por su alto poder de cómputo y por las dimensiones pequeñas del componente lo que era fácil integrarlo, por ejemplo, a un casco militar.

La arquitectura del ODROID XU3 presenta un bloque multimedia soportado por una poderosa GPU. El GPU es un componente de hardware con características multi-core que puede paralelizar cálculos en dicha unidad de proceso gráfico (GPGPU). En la actualidad

muchos móviles modernos poseen GPU en el que se aprovecha el poder procesamiento paralelo.

ODROID XU3 es una nueva generación de dispositivos poderosos, eficientes en el consumo de energía y factor de forma pequeños. Una característica interesante es que este procesador tiene soporte para ejecutar diferentes sistemas operativos tales como Ubuntu y Android versión 4.4.x. y superior. Está equipado con 4 núcleos grandes (Samsung Exynos-5422 ARM Cortex-A15 hasta 2.0 GHz) y 4 núcleos pequeños (ARM Cortex-A7 hasta 1.0 GHz). Posee capacidades de procesamiento mejoradas a bajo consumo de energía. Como parte de la arquitectura tienen una solución denominada big.LITTLE HMP (Heterogeneous Multi-Processing (HMP) solution) lo que permite que los procesadores Exynos-5422 puedan utilizar los 8 cores para administrar computacionalmente las tareas intensivas.

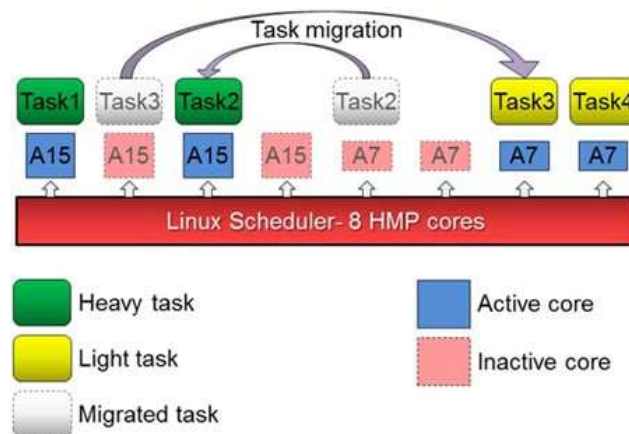


Figura 7.21 – ODROID XU3. Solución Heterogeneous Multi-Processing (HMP) solution (fuente: Hardkernel)

Una de las características principales del ODROID XU3, además de su poder de procesamiento, es que posee un pequeño factor de forma reduciendo de manera considerable el tamaño del dispositivo. Las dimensiones de la placa de circuito impreso son de 94x70x18mm, lo que permite adaptarlo a un casco militar o dispositivo de transporte externo sin ocupar demasiado espacio ni peso.



Figura 7.22 – ODROID XU3
(fuente: Hardkernel)

El ODROID XU3, posee el GPU ARM Mali-T628 MP6 y ofrece soporte para OpenGL ES 1.1, OpenGL ES 2.0 y OpenGL ES 3.0, OpenCL 1.1 como también Google RenderScript. El GPU Mali-T628 está optimizado para procesar aplicaciones gráficas en 3D, computación visual, RA, generación de texturas y reconocimiento de voz, entre otros.



Figura 7.23 – ODROID XU3 implementado en el laboratorio del CITEDEF
(fuente: propia)

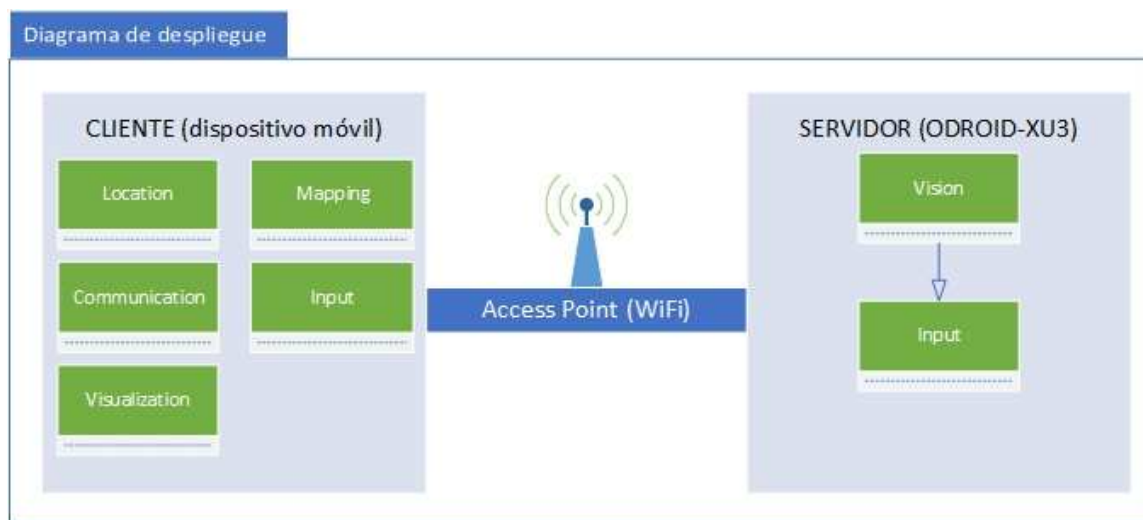
7.5 Proceso distribuido

Como parte de la arquitectura del sistema propuesta y siguiendo los niveles de procesamiento de RA sugeridos por [23], en la vista de despliegue (*deployment*), los elementos presentados en la *Vista de Componentes y Conectores* (figura 7.17) son asignados al hardware de la plataforma donde se ejecuta el sistema (ODROID XU3) del lado del Servidor, es decir, aquellas funcionalidades que incurren en un procesamiento de imágenes intensivo.

El proceso distribuido planteado está soportado por el diseño de una arquitectura Cliente / Servidor.

Del lado del Cliente se ejecutan los procesos menos intensivos tales como:

- Captura de video
- Uso de sensores (ubicación, tracking)
- Comunicación
- Mapeo
- Renderizado de imágenes



**Figura 7.24 – Diagrama de despliegue
(fuente: propia)**

Mientras que del lado del Servidor se implementó el módulo de *Vision* (y algunas funcionalidades del módulo *Input* de entrada de datos) ya que en este módulo se instauró el procesamiento de imágenes pues el ODROID XU3 soporta cargas intensivas de cómputo permitiendo distribuir dicha carga sobre sus procesadores.

Con el objetivo de implementar el procesamiento distribuido se evaluó la comunicación de datos entre un *Cliente* (implementado en JAVA-Android) y un *Servidor* (implementado en JavaScript-Node.js), utilizando Socket.io como tecnología de comunicación entre estos componentes. Cabe aclarar que en el lado del Servidor se ejecutarán los algoritmos pertenecientes al módulo de *Vision*. El funcionamiento entre el Cliente y el Servidor se ejecuta mediante el envío de un *frame* (imagen) desde la aplicación Cliente hacia la

aplicación Servidor. Para lograr esta acción se codifica en formato JSON de la siguiente manera:

```
private void sendImage(Uri uri) throws JSONException, FileNotFoundException {  
  
    JSONObject data = new JSONObject();  
    data.put("image", encodeImage(uri));  
  
    mSocket.emit("new message", data);  
}
```

El servidor recibe el mensaje con la imagen. Guarda el archivo en el *filesystem* y luego crea un proceso donde se ejecuta el algoritmo (implementado en Python) para el procesamiento de la imagen. En este caso se utiliza el algoritmo FAST. El resultado se guarda como una imagen y por otro lado, la respuesta (en formato JSON) es enviada al Cliente.

Los componentes de software instalados en el lado de Cliente fueron Android 6.0, OpenCV 3.1, Socket.io y librerías (ver sección 7.3.2 - *Capa: Librerías*); mientras que en la lado Servidor se instaló Ubuntu Mate versión 16.04, Python versión 3.0, OpenCV 3.1 y Node.js.

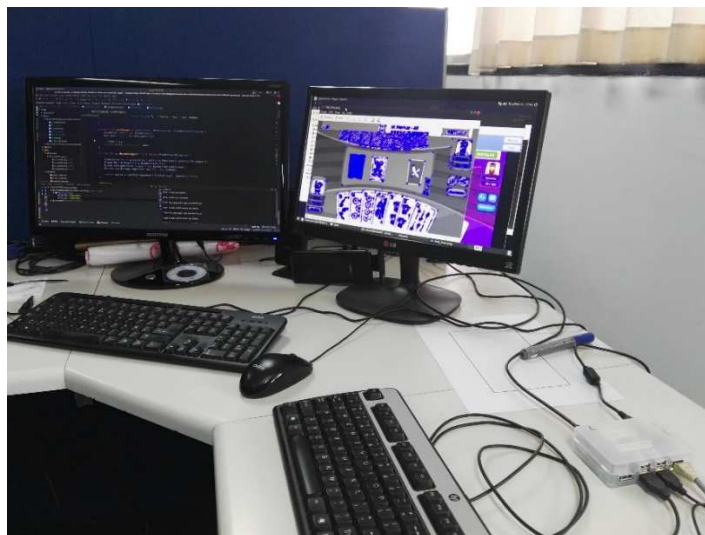


Figura 7.25 – Extrayendo características de una imagen mediante procesamiento distribuido utilizando un ODROID XU3 y una PC que simula a un dispositivo móvil (fuente: propia)

7.6 Resumen del Capítulo

En este Capítulo se estructuró en función de la arquitectura del framework propuesto. Se detalló la metodología de desarrollo sugerida SCRUM+DCU para abordar el diseño del framework y el diseño de las interfaces de las aplicaciones de prueba de concepto. La propuesta de desarrollo está dirigida por la CS, utilizando un método ágil (SCRUM) y centrada en el usuario (DCU). En resumen, la forma de trabajar se centró principalmente en determinar la metodología de identificación y clasificación de requisitos teniendo en cuenta que la problemática principal es la CS. Para ello se analizó la utilización de una metodología en dónde el usuario sea el centro del diseño de las interfaces y de la interpretación y uso de las funcionalidades principales que debería soportar el framework. La utilización de un proceso ágil fue de mucha utilidad pues mediante el desarrollo de prototipos se fueron ajustando las aplicaciones que extendían del framework. Esta combinación de metodologías, es decir, DCU y SCRUM, permitió llegar al prototipo integrado cumpliendo con los requisitos funcionales aportados por los usuarios principales.

Por otra parte, se enfatizan los requisitos no funcionales que si bien no fueron detallados en la presente sección es importante mencionarlos ya que formaron parte de la experiencia de los desarrollos realizados. Si bien la evaluación del framework y de las aplicaciones desarrolladas se constatará empíricamente en las secciones posteriores, es importante mencionar algunos requisitos no funcionales aportados en el presente trabajo, entre ellos se destacan:

- Usabilidad (Efectividad, eficiencia y satisfacción)
- Mantenimiento
- Interoperabilidad entre dispositivos
- Soporte a la CS

Cabe destacar que el presente trabajo se sitúa en el “puente” entre disciplinas como la Ingeniería del Software, mediante el uso de metodologías ágiles, y la Experiencia del Usuario, por medio del uso del DCU, de la que hay distintas publicaciones que tratan el tema tal es el caso de [79] y [80], entre otros.

Se consideró una arquitectura basada en capas a fin de facilitar la integración tecnológica. Se definió la arquitectura del sistema especificando los componentes de hardware que se utilizaron para las pruebas de concepto. Se adaptó la arquitectura de software en función de los requisitos brindados por los usuarios principales logrando cubrir los aspectos de la CS en su totalidad.

Por último, se detalló el proceso distribuido Cliente / Servidor entre los dispositivos móviles y el mini board ODROID XU3 logrando procesar las imágenes en un servidor remoto, móvil, liviano y poderoso, obteniendo muy buenos resultados de respuesta en un ambiente controlado.

Una vez que la arquitectura multicapa ha sido introducida, y gracias a su flexibilidad, es posible ofrecer diferentes despliegues tecnológicos asociados al framework propuesto.

Capítulo 8

Despliegue de la solución

8.1 Introducción

En este capítulo se ha trabajado el despliegue de la solución propuesta. Para ello se han seleccionado e integrado distintas tecnologías, nodos de procesamiento y facilidades de comunicación e interacción necesarias para poder utilizar la tecnología descrita en el capítulo anterior.

Los despliegues propuestos en este capítulo tienen en cuenta las distintas exigencias, tanto funcionales como no funcionales, identificadas en capítulos previos, como también con la CS, la RA y la usabilidad general de los desarrollos y despliegues propuestos en esta tesis doctoral. En este sentido, en este capítulo se remarcan las facilidades logradas y relacionadas con:

- Identificación de la localización del usuario,
- Identificación del entorno (construcciones edilicias, aliados, enemigos, etc.),
- Facilidades de operación, considerando técnicas interactivas basadas en gestos,
- Compatibilidad e interoperabilidad de dispositivos.

El objetivo de este capítulo es demostrar la viabilidad de la solución propuesta y de su distribución.

8.2 Dispositivos utilizados

Con el objetivo en mente de validar la propuesta recogida en esta tesis doctoral, a la hora de realizar el despliegue de la misma se ha trabajado inicialmente en la selección de los dispositivos, las características comunes entre ellos serían nodos de procesamiento móviles, basados en Android y con distintas características. Concretamente, los dispositivos seleccionados se recogen en la tabla 8.1.

En la figura 8.1 se muestran los dispositivos móviles que se han seleccionado para el despliegue. Los cuatro móviles se han usado como dispositivos de adquisición de video, utilización de sensores y visualización de imágenes en la pantalla de los mismos, en el caso de la gafa *Epson Moverio BT 300* la proyección del video se realiza sobre el lente mientras que en los smartphone se utiliza el *cardbox* como adaptador para la proyección y visualización.

Tabla 8.1 Dispositivos móviles. Características técnicas.

Dispositivo	Cámara	SO	CPU	GPU	Conectividad	Sensores
Epson Moverio BT-300	5 Mp	Android 5.1 API Level 22	Intel® Cherry Trail Atom™ x5 1.44GHz Quad Core	Intel HD Graphics (Cherry Trail)	GPS	Geomagnetic sensor/accelerometer/gyroscopic sensor/illumination sensor
Huawei G8	Primaria resolución 13 Mp – 30 fps (a resolución máxima)	Android 5.1.1 Lollipop	Qualcomm Snapdragon 616 MSM8939v2 4 x 1.5 GHz ARM Cortex-A53 4 x 1.2 GHz ARM Cortex-A53	Qualcomm Adreno 405	GPS con soporte A-GPS y GLONASS. 4G/LTE. WiFi (b, g, n). Bluetooth 4.0	Acelerómetro, giróscopo (emulado), proximidad, brújula, luz ambiente, lector de huellas dactilares
Samsung SM-G930F	Principal de 12 megapíxeles con apertura f/1.7, tecnología Dual Pixel Sensor, OIS y dual-LED. Secundaria de 5 megapíxeles con apertura f/1.7	Android 6.0.1 Marshmallow con Samsung TouchWiz	Samsung: 4 núcleos Exynos M1 Mongoose a 2,6/2,3 GHz + 4 núcleos ARM Cortex-A53 a 1,5 GHz (GTS)	Samsung: Mali-T880 MP12 a 650 MHz	4G LTE Cat.12/13, Wi-Fi doble banda, Wi-Fi Direct, Bluetooth 4.2, NFC, GPS con A-GPS y GLONASS, Micro-USB 2.0, USB Host, jack de 3,5 mm	Acelerómetro, giróscopo, proximidad, brújula, luz ambiente, barómetro
ODROID XU3	Externa. Logitech	Android 5.1.1 Lollipop	Samsung Exynos 5422 4 x 2.0 GHz ARM Cortex-A15 4 x 1.4 GHz ARM Cortex-A7 Total de núcleos: 8	Mali-T628 MP6 (OpenGL ES 3.0/2.0/1.1 and OpenCL 1.1 Full profile)	Módulo USB externo GPS. Módulo USB externo WiFi (b, g, n). Módulo USB externo Bluetooth 4.0	No posee



**Figura 8.1 – Dispositivos móviles utilizados para en los despliegues
(Por orden de aparición: Epson Moverio BT 300, Hauwei G8, Samsung SM-G930F,
ODROID XU3)
(fuente: varios)**

8.3 Despliegues considerados

Las arquitecturas hardware utilizadas para realizar el despliegue se centraron en dos tipos, por un lado, *optical see-through* mediante la utilización de las gafas de RA y un mini board ODROID XU3 como servidor para el procesamiento de imágenes y, por el otro, *video see-through* mediante el uso de un *glass cardbox* como soporte a los diferentes smartphone propuestos más un mini board ODROID XU3 con el mismo fin propuesto en la arquitectura anteriormente mencionada.

A continuación, se detallan los dos despliegues que se han probado pero cabe aclarar que únicamente para la evaluación (Capítulo 9) será utilizada la opción 1:

Tabla 8.2 Despliegues utilizados para validar la solución propuesta.

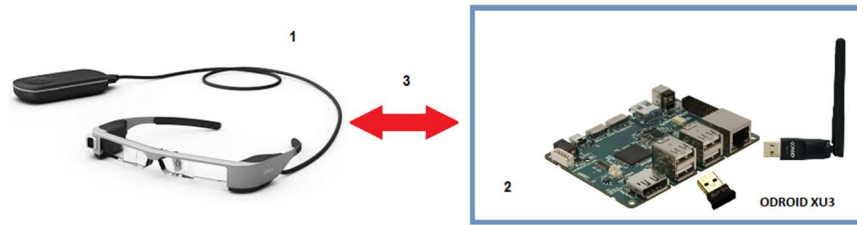
Opción 1	Opción 2
<ul style="list-style-type: none"> • Componentes de hardware: <ul style="list-style-type: none"> ○ <i>Gafa de RA</i> <ul style="list-style-type: none"> ▪ Antena WiFi ▪ Cámara HD ▪ Sensores: Sensores: GPS, acelerómetro, giróscopo, brújula magnética ○ <i>ODROID XU3</i> <ul style="list-style-type: none"> ▪ Antena Bluetooth ▪ Antena WiFi 	<ul style="list-style-type: none"> • Componentes de hardware: <ul style="list-style-type: none"> ○ <i>Glass cardboard</i> ○ <i>Smartphone</i> <ul style="list-style-type: none"> ▪ Antena WiFi ▪ Cámara HD ▪ Sensores: GPS, acelerómetro, giróscopo, brújula magnética ○ <i>ODROID XU3</i> <ul style="list-style-type: none"> ▪ Antena Bluetooth ▪ Antena WiFi
<p>El ciclo de captura de video se procesa de la siguiente manera:</p> <ol style="list-style-type: none"> 1) El video es capturado desde la cámara de la gafa de RA 2) Los frames del video capturado desde la cámara de la gafa de RA son enviados al <i>mini board ODROID XU3</i> <p>Se utilizó el módulo de <i>Communication</i> del framework que utiliza <i>Socket.io</i> como tecnología de comunicación entre la gafa de RA y el <i>mini board ODROID XU3</i>.</p>	<p>El ciclo de captura de video se procesa de la siguiente manera:</p> <ol style="list-style-type: none"> 1) El video es capturado desde la cámara del Smartphone 2) Los frames del video capturado desde la cámara del smartphone son enviados al <i>mini board ODROID XU3</i> <p>Se utilizó el módulo de <i>Communication</i> del framework que utiliza <i>Socket.io</i> como tecnología de comunicación entre el <i>Smartphone</i> y el <i>mini board</i>.</p>

Los smartphone seleccionados se acoplaron al *glass cardboard* emulando la técnica de *video see-through*:

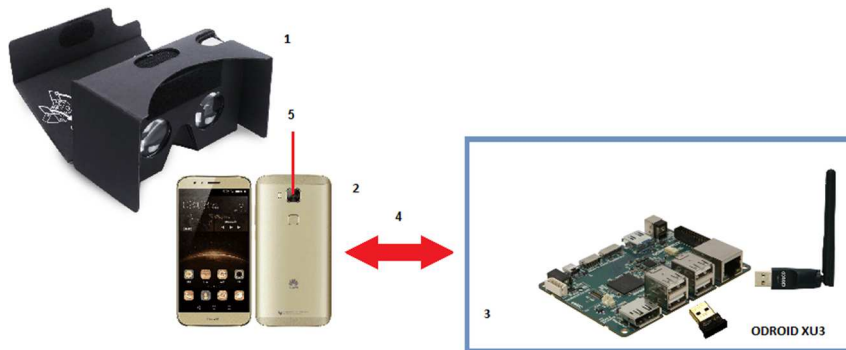


Figura 8.2 – *Glass cardboard* acoplado a los smartphone utilizado en los despliegues (fuente: propia)

A continuación, se ilustran las dos opciones de arquitecturas propuestas:



Opción 1: Optical see-through. 1. Gafa de RA EPSON MOVERIO BT-300; 2. ODROID XU3; 3. Comunicación entre la Gafa de RA y el ODROID XU3



Opción 2: Video see-through. 1. Glass cardboard; 2. Smartphone; 3. ODROID XU3; 4. Comunicación entre los ODROID XU3; 5. Smartphone utilizando la cámara HD.

Figura 8.3 – Opciones de arquitecturas de sistemas propuestos (fuente: varios)

Una vez mencionados los componentes de hardware utilizados para el despliegue de la solución, se está en condiciones de detallar los prototipos evolutivos como parte de la metodología sugerida en la sección 7.2 - *Metodología de desarrollo utilizada* y la implementación de las funcionalidades obtenidas para dar soporte a facilidades relacionadas con la CS y que dará origen a la validación de las aplicaciones desarrolladas utilizando el framework propuesto.

Además de la infraestructura descrita se utilizó un *access point* dentro de un mismo segmento de red (LAN) desde donde se conectaban los dispositivos móviles.

Seguidamente, se analiza, de manera más detenida, algunos aspectos que se consideran fundamentales para el éxito de la solución propuesta. Concretamente se discute el despliegue de los módulos de comunicaciones, de visión y de integración de la tecnología.

8.4 Módulo Communication

Ampliando el detalle de este módulo abordado en la sección 7.3.1.4 - *Módulo Communication*, cabe destacar algunas cuestiones adicionales. Durante su desarrollo y validación se desarrollaron dos aplicaciones extendiendo el framework ejecutándose en el sistema operativo Android. Una aplicación se ejecutaba como un módulo Cliente y la otra aplicación cumplía el rol de Servidor:

- *Módulo móvil Cliente*: se encarga de enviar las imágenes y audio del contexto. Geo posicionamiento y tracking del operador. Adicionalmente, se prueba el protocolo de comunicación entre los módulos.
- *Módulo móvil Servidor*: obtiene las imágenes y audio enviados por el Módulo móvil cliente. Procesa las coordenadas obtenidas por el GPS geo-posicionando al operador.

Con ello se logró integrar el uso de mapas vectoriales (por medio de la librería *mapsforge*), el envío de coordenadas de GPS a través de un socket, el uso de video y audio bidireccional (a través de la librería *libstreaming* utilizando el protocolo *rtsp*). De esta forma, se obtuvo información provista por el GPS como la altitud, velocidad de desplazamiento del operador, longitud y latitud. Por otro lado, se implementó una clase separada donde medía la potencia del *access point* (WiFi). Esta información es fundamental para dar soporte a la CS, el usuario de los despliegues realizados es capaz de consultar y conocer su situación, orientación y dirección en todo momento. Además, gracias al conocimiento de dicha información es posible facilitarle información adicional sobre el entorno en el que se encuentra.



**Figura 8.4 – Módulo Cliente y Módulo Servidor
(fuente: propia)**

Además, con el módulo *Communication* se integraron facilidades interactivas basadas en gesto. Para ello se desarrolló una aplicación que extendía el framework y se integró la librería *OpenCV* en la misma. El objetivo principal es integrar el reconocimiento gestual. Para lograr dar soporte a la interacción gestual se implementaron técnicas como envolvente convexa y tracking de puntos en la palma de la mano. Adicionalmente, al rotar la mano se aplican transformaciones afines y de perspectiva que logran acompañar el movimiento de la mano al detectar los puntos que forman el área o región de interés (ROI) donde se superimprimirá la información sintética correspondiente. En dicha prueba se utilizó una imagen (mapa) del tipo jpg.



Figura 8.5 – Prueba de Reconocimiento gestual (fuente: propia)

8.5 Módulo Vision

Adicionalmente a lo detallado en la sección 7.3.1.6 - *Módulo Vision*, cabe destacar que este módulo es de vital importancia para el framework diseñado. Se integraron los algoritmos de extracción de características en el módulo de *Vision*. Para ello se utilizó la librería *OpenCV* y algunas estrategias de optimización. El algoritmo de detección de características utilizado fue ORB. Se entrenaron cinco construcciones edilicias, de las cuales se extrajeron características las que fueron almacenadas en una base de datos. A medida que se iba recorriendo las instalaciones edilicias, se visualizaba una imagen representativa donde visualmente se daba a conocer al operador de la aplicación que dicho objeto se estaba procesando, tanto el reconocimiento como la clasificación de dicho objeto. Luego la aplicación imprimía en la parte superior izquierda de la pantalla, el texto del objeto reconocido y clasificado.



Figura 8.6 – Prueba de extracción de características utilizando ORB. Reconocimiento y Clasificación de objetos edilicios (fuente: propia)

En el despliegue realizado se integró una aplicación de visualización de alertas y amenazas mediante el uso de iconografía y simbología militar. Para ello se implementó en la aplicación una función de visualización donde se imprime en el campo de visión del operador la información contextual obtenida remotamente desde una unidad de comunicación. El objetivo es potenciar las facultades de la CS.

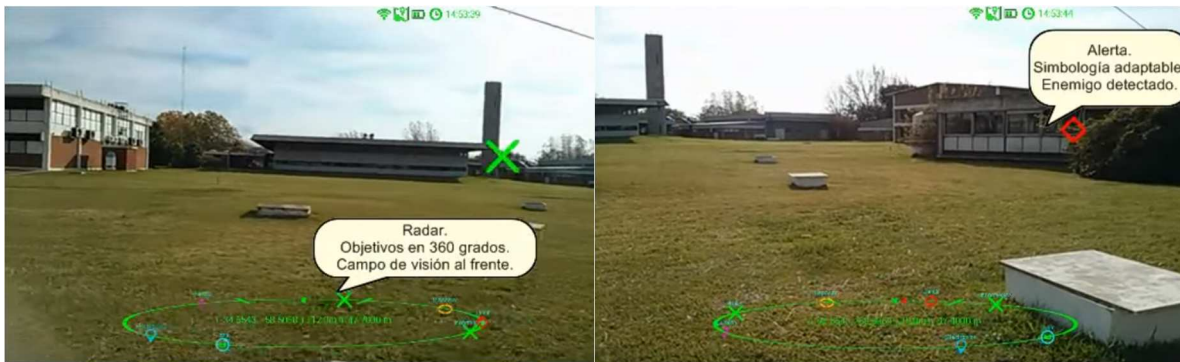
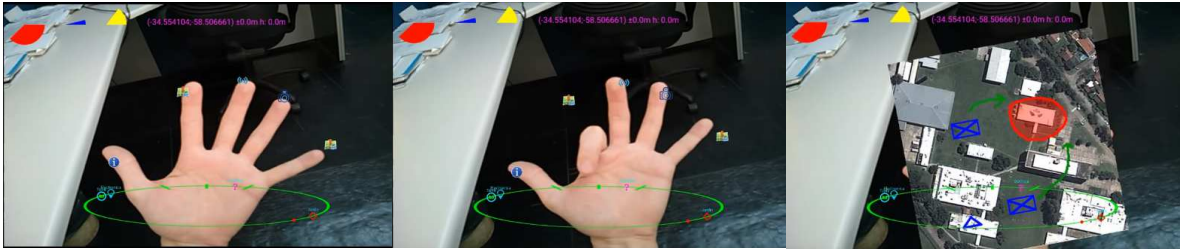


Figura 8.7 – Visualización de iconografía y radar de 360 grados con información contextual impresa (fuente: propia)

Adicionalmente, se integró el framework propuesto en una aplicación para el manejo de menús a través de reconocimiento gestual. Por ejemplo, en la siguiente figura se puede apreciar como el dedo índice de la mano izquierda del operador controla la activación y desactivación de un mapa digital. El objetivo es potenciar la CS y la usabilidad de la solución desarrollada.



**Figura 8.8 – Activación de un mapa por medio de reconocimiento gestual
(fuente: propia)**

8.6 Integración de tecnologías

El despliegue da soporte a las necesidades funcionales y no funcionales identificadas. Para ello se integró una aplicación software para que sea ejecutada en cada uno de los dispositivos propuestos utilizando la arquitectura de sistema mencionada en la sección 8.3 - *Despliegues*.

Entre las funcionalidades desarrolladas en la aplicación y extendidas del framework se implementaron las siguientes características:

- *Visualización de objetivos tácticos en un radar en 360 grados*: los objetivos tácticos son cargados en la aplicación para que sean visualizados en la misma mediante la iconografía, que, por defecto, se diseñó. Los objetivos representan coordenadas latitud y longitud que son interpretados por el paquete *Location* del framework. Luego, los símbolos tácticos son impresos en el radar en 360 grados, permitiendo al operador tener una “visión” perimetral de las amenazas u objetivos tácticos.
- *Posicionamiento y orientación espacial de los objetos y del operador*: el norte magnético sirve de referencia para ubicar y orientar al operador. Mientras el operador se desplaza, se calcula su posición relativa con respecto al norte magnético. De la misma manera que en la funcionalidad anterior, se utiliza el paquete *Location* a fin de calcular la ubicación y orientación del usuario.
- *Ayudas visuales mediante iconografía/simbología adaptable*: la iconografía o símbolos tácticos utilizados en el radar de 360 grados, son adaptables y configurables permitiendo identificar la/s amenaza/s u objetivo/s táctico/s de acuerdo a la simbología militar, en este proyecto en particular. No obstante, se puede utilizar otra simbología, como, por ejemplo, iconos logísticos u otros que sirvan de referencia para el operador.

- *Uso de menús interactivos mediante Reconocimiento Gestual:* mediante la utilización de técnicas de reconocimiento gestual se puede obtener información a demanda. Esta funcionalidad se implementó en el paquete *Input* del framework. Si bien esta funcionalidad posee características de procesamiento de imágenes y visión por computador, no fue implementado en el paquete de *Vision* ya que se analizó dentro de la lógica del framework que era una mejor opción de diseño implementarlo en el paquete *Input*.
- *Reconocimientos de objetos edificios:* como parte principal de esta prueba de concepto se implementó en la aplicación desarrollada la posibilidad de reconocer varios objetos edificios previamente entrenados. Esta funcionalidad fue implementada en el módulo de *Vision*. No obstante, para reconocer el primer frame del objeto a clasificar, se utilizó la técnica de reconocimiento gestual donde, a través de la detección del dedo anular y del dedo índice del operador, se especifica el área aproximada de captura del primer frame. A partir de esta detección inicial, el siguiente paso es la clasificación del objeto en cuestión utilizando las técnicas de reconocimiento de objetos adecuada implementadas en el framework y extendida en la aplicación de prueba.
- *Uso de cartografía en 3D mediante uso de RA:* esta funcionalidad se encuentra implementada en el módulo de *Mapping*. Este módulo implementa una interfaz a un SIG completamente funcional que permite cargar cartas prediseñadas a la aplicación y ubicar elementos para su posterior visualización y manipulación ya sea en una carta 2D como en un mundo 3D.

Básicamente se integraron en una aplicación todas las funcionalidades demostradas en cada uno de los prototipos desarrollados anteriormente.

En las siguientes figuras se puede observar cada una de las características implementadas en la aplicación utilizando el framework propuesto:



a) Radar 360 grados

b) Reconocimiento de objeto (Edificio)



c) Reconocimiento gestual e iconografía

d) Visualización de mapa activado por gesto

Figura 8.9 – QR para acceder al vídeo en el cual se muestra la integración de las funcionalidades principales del framework (fuente: propia)



<https://youtu.be/hCSYqWnIs48>

8.7 Resumen del Capítulo

En este capítulo se ha detallado el despliegue de la solución propuesta en esta tesis doctoral teniendo en cuenta las consideraciones descritas en los capítulos previos, especialmente en los capítulos quinto y sexto. El despliegue realizado integra características relacionadas con

la CS, la RA y la usabilidad del desarrollo. Para ello, en este capítulo, se han integrado componentes con capacidades para soportar la localización del usuario, la comprensión del entorno y, también, se han potenciado las facilidades interactivas activas y pasivas del usuario del sistema desplegado.

Los despliegues realizados serán la base para la actividad de validación que se recogerá en el próximo capítulo.

Capítulo 9

Evaluación

9.1 Introducción

Una vez descrito el framework propuesto, en sus distintas dimensiones –tanto teóricas como técnica-, cabe abordar actividades de evaluación.

La provisión de métodos válidos y confiables para evaluar la Conciencia Situacional (CS) es esencial durante el diseño y la evaluación del sistema. Uno de los principales objetivos de cualquier esfuerzo de diseño del sistema es mejorar la CS del personal involucrado en la operación del sistema. Con el fin de garantizar que la CS sea mejorada y no degradada por el esfuerzo de diseño, se requieren técnicas válidas y confiables que puedan medir con precisión la CS del operador y del equipo a lo largo del proceso de diseño. Las técnicas de medición / evaluación más conocidas de la CS son: SAGAT [81] [82], SART [81] y SPAM [81] [83].

La técnica SAGAT se utiliza principalmente durante las simulaciones. En este caso, la simulación está congelada y el/los participante/s debe/n responder algunas preguntas. El resultado de esta técnica es el nivel de comprensión del participante en un entorno realista.

La Técnica SART es una escala de autoevaluación del 1 al 7, en la cual se les pide a los participantes que califiquen su comprensión.

El Método SPAM se basa en la interacción entre la máquina y el ser humano.

Para llevar a cabo las actividades de evaluación de las implementaciones asociadas con esta Tesis Doctoral, y presentadas en el capítulo anterior, fueron utilizadas las técnicas SAGAT y SART para evaluar la CS en entornos hostiles. La primera técnica se utilizó para evaluar una situación simulada en un entorno real utilizando una encuesta de elaboración propia desarrollada para tal fin. En cuanto a la segunda técnica fue la más adecuada para establecer un valor cuantitativo de referencia para la evaluación en una escala de 5 puntos.

Las técnicas anteriores fueron adaptadas y complementadas mediante el despliegue de *User Testing* y *Thinking Aloud protocol*. La documentación de las actividades de evaluación se realizó teniendo en cuenta la norma ISO/IEC 25062 (2006).

El objetivo central de la evaluación es obtener información con la que cubrir las hipótesis que se plantean en la evaluación del framework desarrollado y de algunas de sus implementaciones. Las principales hipótesis que se pretenden contrastar en esta evaluación están relacionadas con la usabilidad del despliegue. Concretamente las hipótesis a considerar son las siguientes:

- H_0 : La implementación realizada asegura niveles altos de CS de los participantes en el experimento de evaluación, ya que permite a los usuarios percibir, comprender y, en su caso, tomar decisiones adecuadas con efectividad y eficiencia en un entorno hostil simulado.
- H_1 : La implementación realizada asegura niveles altos de CS de los participantes en el experimento de evaluación, ya que permite a los usuarios percibir, comprender y tomar decisiones adecuadas con alto grado de experiencia de usuario (UX).

Como se ha mencionado la evaluación gira entorno a la usabilidad del despliegue. En este sentido, se evaluarán las facilidades para que los participantes en el experimento logren sus objetivos (efectividad) y lo hagan en tiempo y forma (eficiencia). En lugar de la satisfacción, criterio habitualmente considerado en las evaluaciones de usabilidad, en la H_1 se ha considerado más interesante dar cuenta de la experiencia de usuario experimentada por los participantes en la evaluación. De esta forma se puede valorar de una forma más adecuada criterios de calidad que van más allá de la mera satisfacción que tienen que ver con la utilidad, la accesibilidad, la credibilidad o lo valioso que puede ser el despliegue evaluado por parte de usuarios específicos haciendo actividades concretas en un contexto de uso también específico, como es el militar.

Por lo tanto, en función de las hipótesis evaluaremos el nivel de CS que aporta la implementación desarrollada y presentada en el capítulo anterior.

Partiendo de la información anterior y teniendo en cuenta la norma ISO/IEC 25062:2006 se documentan, seguidamente, las actividades de evaluación llevadas a cabo.

9.2 Resumen de la evaluación

RAIOM (Realidad Aumentada para la Identificación de Objetivos Militares) es el nombre del framework de RA enfocado en mejorar la CS, que se propone en este trabajo de tesis. Con el objetivo de medir su efectividad, eficacia y usabilidad orientada a cumplir con los 3 niveles del modelo propuesto por Endsley con respecto a la CS, es decir, percepción, comprensión y proyección, se evaluó el despliegue para medir el cumplimiento de la CS. El experimento fue realizado por 6 participantes, 3 usuarios expertos con poder de decisión en las operaciones militares y 3 participantes con perfil operativo. Se conformaron dos grupos para tal fin, el Grupo A, estaba formado por los primeros 3 participantes con poder de decisión y el Grupo B que estaba conformado por 3 participantes operativos. Para tal fin se utilizaron las técnicas *User Testing*, *Thinking Aloud* y/o *SAGAT / SART*. Por otro lado, se evaluó la experiencia de usuario de los participantes utilizando como elemento de evaluación el cuestionario UMUX [84]. Los resultados obtenidos en el cumplimiento de la CS evaluando los 3 niveles propuestos por Endsley para el Grupo A, el conjunto de individuos decisores, representa un 88,89%, mientras que para los usuarios del Grupo B representa un cumplimiento del 85,83%. En cuanto a la experiencia del usuario se llega a la conclusión que la interfaz diseñada cumple en un 88,33% para los usuarios pertenecientes al Grupo A y un 86,67% para los usuarios del Grupo B.

9.3 Descripción del producto software

RAIOM (Realidad Aumentada para la Identificación de Objetivos Militares) es el nombre del framework de RA enfocado en mejorar la CS. El diseño de dicho framework consistió en el diseño, principalmente, de la arquitectura de la capa intermedia, es decir, el *middleware*. El framework soporta el filtrado de información prioritaria lo que permite desarrollar aplicaciones cuya información contextual sea accedida por al menos 2 tipos de usuarios diferentes (Filtro de información 1 y 2). El framework soporta procesamiento distribuido lo que garantiza que el procesamiento de imágenes sea realizado en una unidad de procesamiento externa independiente del dispositivo móvil que se utilice. RAIOM se desarrolló y desplegó seleccionando e integrando distintas tecnologías. Concretamente, los

despliegues se implementaron en gafas de RA (tecnología que formó parte de las evaluaciones) y en el ODROID-XU3, y en el glass cardboard (se utilizó a modo comparativo en cuanto a la implementación de la aplicación desarrollada pero no formó parte de la evaluación).

9.4 Objetivos de la Evaluación

Las pruebas fueron realizadas para evaluar un conjunto de características del framework propuesto en los despliegues que se han descrito con anterioridad. Dichas características están relacionadas con evaluar la CS utilizando los tres niveles propuestos por Endsley: percepción (información y elementos del contexto), comprensión (entender el significado y relevancia de la situación) y proyección (futuros estados y eventos) y, adicionalmente, se analizó la experiencia del usuario en utilizar la interfaz desarrollada. Dichos niveles se evaluarán a través de la consideración de los niveles de efectividad, eficiencia y experiencia del usuario con que usuarios específicos en contextos de uso también específico podían llevar a cabo tareas relacionadas con la percepción, comprensión y, en su caso, toma de decisiones en un entorno hostil.

9.5 Método utilizado para la evaluación

9.5.1 Participantes

En la evaluación fueron seleccionados participantes representativos para formar dos grupos. Por un lado (Grupo A), 3 usuarios Jefes de Comandos (decisores) y, por otro (Grupos B), 3 usuarios comandos (operativos). Todos los participantes tenían las siguientes características:

- a) Formación militar en comandos
- b) Conocimiento de la simbología militar
- c) Familiarización con qué es la CS en la jerga militar
- d) Experiencia mayor a 3 años en funciones de comandos

Los participantes fueron seleccionados en el CITEDEF (Instituto de Investigaciones Científicas y Técnicas para la Defensa) y en el DIC 601, Campo de Mayo, Argentina.

Tabla 9.1 Perfil de los participantes.

Participante	Grupo	Establecimiento	Años de experiencia	Edad	Sexo	Nivel de uso de tecnología (alto/medio/bajo)
1	A	DIC 601	>15	55	M	Medio
2	A	DIC 601	>15	52	M	Medio
3	A	CITEDEF	>10	40	M	Medio
4	B	DIC 601	>10	44	F	Medio
5	B	CITEDEF	>10	32	M	Alto
6	B	CITEDEF	>10	43	M	Alto

9.5.2 Tareas

Descripción de escenario: a todos los participantes se les explicó verbal e individualmente el escenario creado a fin de evaluar las aplicaciones desarrolladas. En este escenario, se recreó una operación donde el objetivo principal estaba a 500 metros de distancia del grupo comando. El contexto simulado se trataba de una situación hostil y de alto riesgo. La misión consistía en acercarse lo más posible al objetivo, sin ser detectados, e informar de la situación. El sistema era utilizado por el JPat (Grupo A, decisor) y por una Pat (Grupo B, operativo). Por lo tanto, las tareas que debían ejecutar el Grupo A estuvieron relacionadas con los 3 niveles de la CS propuestos por Endsley, es decir, percepción, comprensión y proyección. Mientras el Grupo B estuvo relacionadas con los niveles de percepción y comprensión. Al Grupo B no se le plantearon tareas de toma de decisiones, por lo tanto, no se evaluó el nivel de proyección. En cuanto al Grupo A, la actividad más importante estuvo relacionada con la toma de decisiones, pues es necesario contar con:

- una correcta *percepción* de los elementos que conforman cada situación,
- una adecuada *comprensión* de la situación actual, y
- una clara *proyección* de cuáles son los posibles escenarios futuros fruto de las decisiones tomadas.

Para que los participantes lograsen llevar a cabo el escenario antes descrito es necesario que dispongan de claves/información que facilite la conciencia. En la siguiente Tabla se identifican las tareas concretas que se les pidió realizar a los participantes en el experimento. A los participantes en el experimento se les pedía que analizarán el contexto y que, en función de dicho análisis, fueran capaces de determinar su ubicación (dónde estaban), y que

obtuviesen información del mismo (qué y quién estaba en esa ubicación). A los participantes se les hacían preguntas concretas relacionadas con esa información y con ello fue posible evaluar la eficiencia de la percepción. En el caso de la comprensión se pedía a los participantes que describieran el contexto atendiendo a la identificación de amenazas y a la situación (ubicación y distancia de las mismas).

Finalmente, y solo para los participantes del Grupo A, se solicitó a los participantes que tomaran decisiones relacionadas con su seguridad y la de sus comandos. Es decir, los participantes del Grupo A deben organizar y salvaguardar los recursos humanos y materiales que dependen de ellos.

De acuerdo a lo anteriormente descrito se identificaron las tareas que debían realizar los participantes involucrados en el experimento de evaluación:

Tabla 9.2 Tareas que debían realizar los participantes.

Nivel	Grupo	Tareas
Percepción	A y B	Determinar la ubicación espacial Obtener información del entorno
Comprensión	A y B	Analizar la peligrosidad de las amenazas Determinar la ubicación y distancia hacia el objetivo
Proyección	A	Proyectar vías de escape Organizar recursos humanos y materiales en caso de escape

Para identificar y proponer las tareas que cada grupo debía realizar se utilizó la taxonomía propuesta por Bloom [85] para elaborar las preguntas que fueron hechas a cada uno de los participantes del test con usuarios. Estas preguntas estaban relacionadas con las tareas que debían ejecutar de acuerdo a los tres niveles de la CS sugeridos por Endsley y considerados en esta evaluación.

En la siguiente Tabla se da cuenta de las acciones (Bloom) y preguntas formuladas a los participantes para valorar su nivel de efectividad y eficiencia en la percepción, comprensión y proyección.

Para valorar la efectividad se tabularon las respuestas correctas a las preguntas realizadas. Para ponderar la eficiencia, además, se tomaron tiempos invertidos por los usuarios en facilitar las respuestas a las preguntas formuladas.

Tabla 9.3 Preguntas elaboradas para comprobar el grado de efectividad y eficiencia con el que los participantes percibían, comprendían o proyectaban en un contexto dado.

Proc. Cognitivo 3 Niveles	Acciones	Preguntas
Percepción	Describir Encontrar Localizar Identificar Reconocer	Q1: ¿Cuál es tu ubicación geográfica? ¿La puedes establecer? Q2: ¿Puedes reconocer esa construcción? Q3: ¿Puedes encontrar el objetivo de la misión? Q4: ¿Puedes identificar las amenazas?
Comprensión	Clasificar Interpretar Inferir Explicar	Q5: ¿Podrías explicar que está ocurriendo a tu alrededor? Q6: ¿Podrías clasificar el tipo de amenaza? Q7: ¿Qué puede decir sobre la distancia hacia el objetivo? ¿Tienes esa información? Q8: ¿Podrías afirmar que la información disponible es suficiente para determinar la peligrosidad del entorno?
Proyección	Calcular Usar Realizar Planear	Q9: ¿Cómo usarías la información provista por el sistema para preservar los medios humanos y materiales? Q10: ¿Podrías planificar una vía/camino de escape? (teniendo en cuenta las amenazas circundantes) Q11: ¿Cómo organizarías tus medios humanos y materiales en caso de ataque?

Reconocimiento vocal: con el objetivo de ejecutar acciones determinadas del sistema y también con la necesidad de confirmar aspectos de percepción, comprensión y proyección, los participantes podían realizar comandos por voz (Thinking Aloud Protocol). A continuación, se listan las palabras claves que se sugería utilizar a los participantes durante el experimento:

Tabla 9.4 Comandos de voz y Acción.

Comando	Acción
Activar audio	Se activa el servicio de reconocimiento vocal. Por cada acción que se deba ejecutar (comando por voz) se deberá pronunciar el comando “Activar audio”
Encender	Se activa el modo pantalla. La cámara del dispositivo móvil se activa. La función principal es el reconocimiento de objetos e impresión de información contextual.
Amenazas	Una vez activada la pantalla (comando Encender) y ejecutado el comando “Amenazas”, se listan las amenazas circundantes adicionalmente a la iconografía que se visualiza en el radar 360 grados.
Objetivo	Una vez activada la pantalla (comando Encender) y ejecutado el comando “Objetivo”, se detalla el objetivo de la misión adicionalmente a la iconografía que se visualiza en el radar 360 grados.
Apagar	Retorno al modo de pantalla translucido para una mejor visualización e interpretación de la información contextual aportado por el radar 360 grados.

Por último, en cuanto a las instrucciones de tareas a realizar por los participantes se puede consultar el Apéndice B para una lista completa de las tareas realizadas por cada uno de los grupos de participantes.

9.5.3 Ejecución de las pruebas

Las pruebas para evaluar la CS se basan en plantear tareas a los usuarios y que estos respondan a una serie de preguntas seleccionadas de acuerdo a los tres niveles que Endsley propone. Los participantes darán cuenta, respondiendo a las preguntas realizadas, teniendo en cuenta el grado de efectividad y eficiencia con el que pueden dar respuesta a cada pregunta. Las pruebas de evaluación se realizaron en el campo a cielo abierto con el objetivo principal de tener una correcta señal de GPS, una óptima visibilidad y simular una situación lo más real posible. La distancia al hipotético objetivo fue de unos 500 metros desde el punto de inicio de las tareas previstas.

9.5.4 Tecnología utilizada

En las pruebas realizadas los participantes utilizaron las gafas de RA Epson Moverio BT-300, *optical see-through*, y un mini board ODROID-XU3 como servidor para el procesamiento de imágenes. Para simular el bloque CII se implementó un Access Point (Wireless) LinkSys en el que los participantes tenían acceso a través de los dispositivos móviles (gafas de RA y ODROID).

9.5.5 Herramientas del Administrador de prueba

Como se mencionó anteriormente se utilizó, por un lado, un cuestionario para que los participantes pudieran evaluar la efectividad y eficiencia a la hora de percibir, comprender o tomar decisiones en cumplimiento de la CS basado en los 3 niveles sugeridos por Endsley y, por el otro, al final del experimento, se usó un cuestionario UMUX para evaluar la experiencia del usuario (ver el Apéndice A). Los cuestionarios fueron presentados mediante la utilización de preguntas con respuestas basadas en escala Likert de 5 puntos (SART). El proceso fue verbal entre el administrador/coordinador de la evaluación y los participantes en la prueba, centrándose en la técnica SAGAT que propone la parada de la tarea que se está ejecutando para realizar y evaluar las preguntas correspondientes.

9.5.6 Diseño experimental

Se seleccionó 2 niveles de participantes, aquellos relacionados con JPat capacitados para tomar decisiones y aquellos que son militares pero su función es netamente operativa (Pat). El primer grupo, (Grupo A), está compuesto por militares cuyo objetivo es tomar decisiones en el terreno de operaciones en función de la información contextual que obtiene. El segundo grupo de participantes, (Grupo B), son militares cuya finalidad es realizar operaciones en función de la información recibida del JPat y de la información contextual que reciben. En total se seleccionaron 6 participantes, es decir, 3 participantes para el Grupo A y 3 participantes para el Grupo B. Estos grupos realizaron actividades relacionadas con la CS bajo dos premisas centrales, la primera, enfocada en los tres niveles propuestos por Endsley (percepción, comprensión y proyección para el Grupo A y percepción y comprensión para el

Grupo B) mientras que la segunda, se tomó como referencia la taxonomía de Bloom que identifica los procesos cognitivos en diferentes niveles (inferior y superior), concretamente en el nivel inferior, es decir, aquellos relacionados con recordar, comprender y aplicar.

← ———— Procesos cognitivos de orden inferior ———— →

RECORDAR			COMPRENDER			APLICAR		
Recordar hechos/datos sin necesidad de entender. Se muestra material aprendido previamente mediante el recuerdo de términos, conceptos básicos y respuestas.			Mostrar entendimiento a la hora de encontrar información del texto. Se demuestra comprensión básica de hechos e ideas.			Usar en una nueva situación. Resolver problemas mediante la aplicación de conocimiento, hechos o técnicas previamente adquiridas en una manera diferente.		
PALABRAS CLAVE:			PALABRAS CLAVE:			PALABRAS CLAVE:		
Elegir	observar	mostrar	Preguntar	esquematizar	Actuar	emplear	practicar	
Copiar	omitir	deletrear	Generalizar	predecir	Identificar	seleccionar	agrupar	
Definir	rastrear	afirmar	Clasificar	dar ejemplos	Calcular	elegir	resumir	
Decir	cuándo	duplicar	Comparar	relacionar	Entrevistar	planear	desarrollar	
Citar	repetir	qué	Contrastar	ilustrar	Enseñar	transferir	interpretar	
Leer	relacionar	nombrar	Parfrasear	demostrar	Usar	demostrar	categorizar	
Quién	listar	repetir	Informar	discutir	Conectar	dramatizar	construir	
Recitar	escribir	localizar	Inferir	revisar	Planear	manipular	resolver	
Cómo	dónde	Memorizar	Interpretar	mostrar	Simular	seleccionar	unir	
Por qué	reconocer		Explicar	resumir	Hacer uso		organizar	
Traducir			Expresar	observar				
ACCIONES	RESULTADO		ACCIONES	RESULTADO		ACCIONES	RESULTADO	
Describir	Definición		Clasificar	Colección		Desempeñar	Demostración	
Encontrar	Hechos		Comparar	Ejemplos		Ejecutar	Diario	
Identificar	Etiquetado		Ejemplificar	Explicación		Implementar	Ilustraciones	
Listar	Listado		Explicar	Etiquetado		Usar	Entrevista	
Localizar	Cuestionario		Inferir	Listado		Emplear	interpretación	
Nombrar	Reproducción		Interpretar	Esquema		Realizar	Simulación	
Reconocer	Test		Parfrasear	Cuestionario			Presentación	
Recuperar	Cuaderno		Resumir	Resumen			Dibujo	
	Fotocopia			Muestra y cuenta				
PREGUNTAS			PREGUNTAS			PREGUNTAS		
¿Puedes enumerar...?			¿Puedes explicar que está ocurriendo...?			¿Cómo usarías...?		
¿Puedes recordar...?			¿Cómo clasificarías...?			¿Qué ejemplos sobre...puedes encontrar?		
¿Puedes seleccionar...?			¿Cómo compararía/contrastaría...?			¿Cómo organizarías... para presentar...?		
¿Cómo ocurrió...?			¿Cómo podrías parfrasear el significado de...?			¿Cómo aplicarías lo que has aprendido para desarrollar...?		
¿Cómo es...?			¿Cómo resumirías...?			¿Qué enfoque usarías para...?		
¿Cómo describirías...?			¿Qué puedes decir sobre...?			¿Qué aspectos seleccionarías para mostrar...?		
¿Podrías explicar...?			¿Cuál es la mejor respuesta...?			¿Qué preguntas harías en una entrevista a...?		
¿Cómo mostrarías...?			¿Qué afirmaciones apoyan...?					
¿Qué es...?			¿Podrías afirmar o interpretar en tus propias palabras...?					
¿Cuál...?								
¿Quién fue...?								
¿Quiénes fueron los principales...?								
¿Por qué...?								

Figura 9.1 – Taxonomía de Bloom
(fuente: Bloom B. [85])

Como resultados de estas dos premisas permitió describir las actividades y hacer las preguntas a los dos Grupos de usuarios (cuadro 1 – Apéndice A). Al finalizar el experimento los Grupos respondieron un cuestionario relacionado con la experiencia de usuarios (UMUX) (cuadro 2 – Apéndice A)

9.5.7 Procedimiento

A cada participante se le manifestó que el objetivo de la evaluación del despliegue estaba centrado en analizar el cumplimiento de la CS en situación de riesgos en un terreno hostil. Adicionalmente se les informó que dicha evaluación tenía también por objetivo medir la experiencia del usuario. También se les aclaró a todos los participantes que en ningún caso se estaba evaluando sus capacidades o aptitudes militares, sino que la única evaluación realizada estaba relacionada con los artefactos desarrollados y sus posibilidades para dar soporte a actividades cotidianas en el ámbito militar. A cada participante se le explicó cada uno de los conceptos analizados en el experimento y se los entrenó en el uso de la aplicación desarrollado. Cada participante realizó la prueba utilizando el despliegue con unas *gafas de RA* mientras que no se evaluó el despliegue con el glass cardboard acoplado a un smatphone por razones de usabilidad pues resultó más cercano a la realidad analizar el comportamiento con las gafas de RA. No obstante, el despliegue con el glass cardboard fue utilizado por los individuos, a modo comparativo, pero no fue evaluado el uso del mismo. Al inicio del experimento a cada participante se le solicitaba sus datos a fin de asociarlo al grupo de pertenencia (Grupo A o Grupo B). Posteriormente, comenzaban a realizar la prueba guiado por el Administrador del experimento. El experimento se centró en la metodología propuesta por SAGAT / SART y se complementó con el protocolo Thinking Aloud, es decir, a pedido del Administrador, se solicitaba a los participantes que verbalizaran sus respuestas para evaluar la CS por niveles (cuadro 1 – Apéndice A) y, por último, el cuestionario correspondiente a analizar la experiencia del usuario (cuadro 2 – Apéndice A). En el caso del cuestionario por niveles se utilizó SART, presentado con una escala Likert de 5 puntos. De igual forma para el cuestionario UMUX se utilizó una escala Likert también de 5 puntos.

9.6 Presentación y discusión de resultados

9.6.1 Prueba 1 realizada el 10 de noviembre de 2018

Todos los participantes fueron capaces de completar las tareas propuestas con efectividad y eficiencia. Cada uno de los seis participantes realizaron las actividades de manera individual

y posteriormente, los resultados fueron analizados en función del grupo al que correspondían. De acuerdo al grupo de pertenencia, se aplicó el filtro de información prioritaria en la aplicación de prueba; en función de esto se obtuvieron los siguientes resultados:

Tabla 9.5 Resultados de la evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley

Nivel Grupo	Percepción	Comprensión	Proyección	Total
A	4,67	4,33	4,33	4,44
B	4,50	4,08	N/A	4,29
Total	4,58	4,21	4,33	4,38

Tabla 9.6 Resultados de la evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX

Grupo	Total
A	4,42
B	4,33

Con estos resultados se puede ser razonablemente optimista con el grado de consecución de la CS obtenido por los participantes en el experimento, utilizando el equipo y realizando las tareas propuestas con efectividad y eficiencia. Los participantes que pertenecen al Grupo A, el conjunto de individuos decisores, representa un 88,89%, mientras que para los usuarios del Grupo B representa un cumplimiento del 85,83%. Los participantes logran recopilar suficiente información utilizando el equipo facilitado para responder a las preguntas realizadas. Las facilidades de percepción son superiores a las logradas para otros niveles de Endsley (comprensión y proyección). Pero esos otros niveles de Endsley también se responden con un grado alto de efectividad.

En cuanto a la experiencia del usuario (UX) se llega a la conclusión que la interfaz diseñada cumple en un 88,33% para los usuarios pertenecientes al Grupo A y un 86,67% para los usuarios del Grupo B. Por ello se puede manifestar que los participantes se sienten

satisfechos con la implementación del sistema RAIOM facilitada en el experimento y que el mismo les parece útil.

Obviamente, el experimento de evaluación presenta algunas limitaciones en cuanto a su validez. Estas limitaciones se centran en los siguientes aspectos:

- En función del número de participantes involucrados. Esta evaluación debería ser extendida a más participantes para que, estadísticamente, los resultados fueran más concluyentes.
- En función de la recreación realizada y el hecho de que los participantes fueran conscientes de no estar ante una situación hostil real sino simulada al hacer el experimento, conlleva que se desvirtúe el principal escenario en el que el producto implementado debiera presentar su principal utilidad.

Seguidamente se recogen los datos recogidos con la realización del experimento y se da cuenta, gráficamente de dichos datos.

Tabla 9.7 Cuadro de Evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley. Grupo A.

Nivel	Pregunta	Puntaje por participante			
		P1	P2	P3	Total
Percepción	Q1	5	5	5	
	Q2	5	4	5	
	Q3	5	5	5	
	Q4	5	5	2	
		5	4,75	4,25	4,67
Comprensión	Q5	5	4	4	
	Q6	4	4	5	
	Q7	5	5	5	
	Q8	3	4	4	
		4,25	4,25	4,5	4,33
Proyección	Q9	5	5	5	
	Q10	4	4	4	
	Q11	4	3	5	
		4,33	4	4,67	4,33

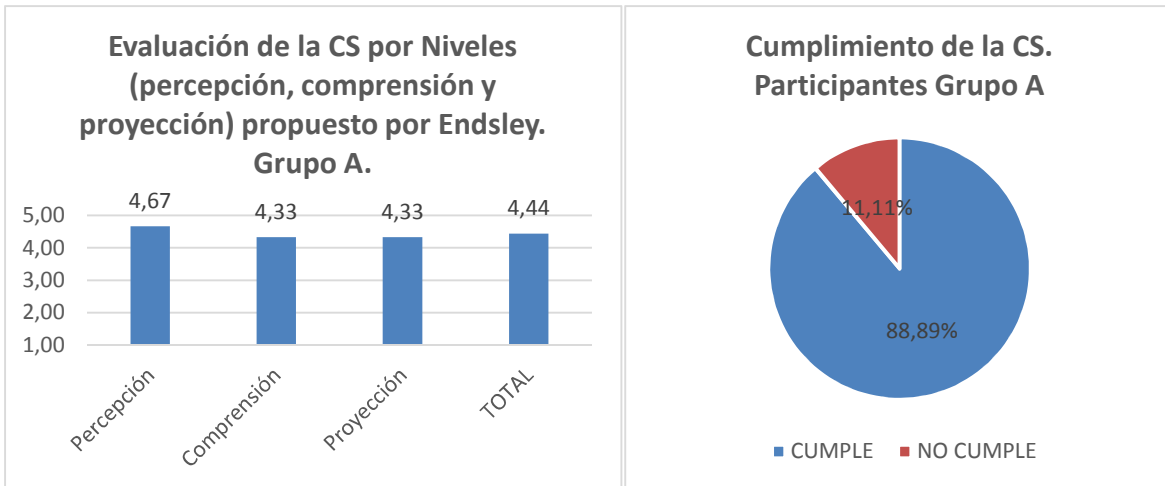


Figura 9.2 – Grafico de Evaluación de la CS por Niveles según Endsley: percepción, comprensión y proyección del Grupo A de participantes evaluados. Gráfico de cumplimiento de la CS de los participantes del Grupo A. (fuente: Propia)

En la figura anterior se muestra las valoraciones de los participantes sobre la percepción, comprensión, proyección y el valor acumulado de esas valoraciones. Se alcanza un valor medio acumulado de 4,44 sobre 5, donde el mayor valor corresponde con el nivel de percepción ofrecido por el sistema desplegado (4.67). De ello se desprende que el nivel de efectividad es alto y que el sistema facilita la CS de los participantes en el experimento. Además, los participantes del grupo A de forma mayoritaria (89%) valoraron positivamente el cumplimiento de sus cometidos por parte del sistema evaluado.

De manera similar, en la Tabla 9.8 y en la Figura 9.3 se da cuenta de los resultados obtenidos por parte de los participantes del grupo B en el experimento de evaluación. Como se ha mencionado los participantes del grupo B no necesitan tomar decisiones. Por ello, a estos participantes solo se le formularon preguntas y se les pidieron valoraciones de la información facilitada por el sistema relacionado con la percepción y la comprensión. Los resultados, aunque buenos, no alcanzaron los niveles de los participantes del grupo A. El valor acumulado de valoración por parte de los participantes del grupo B fue de 4,08

Tabla 9.8 Cuadro de Evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley. Grupo B.

Nivel	Pregunta	Puntaje por participante			
		P1	P2	P3	Total
Percepción	Q1	4	4	5	
	Q2	5	5	4	
	Q3	5	5	5	
	Q4	4	5	3	
		4,5	4,75	4,25	
Comprensión	Q5	4	3	5	
	Q6	4	4	5	
	Q7	5	4	5	
	Q8	4	2	4	
		4,25	3,25	4,75	

Respecto al grado de cumplimiento de la CS por parte de los participantes del Grupo B el 86% de los participantes valoraron positivamente y opinaron que el sistema evaluado permitía cubrir los objetivos de percepción y comprensión del contexto recreado en la evaluación.

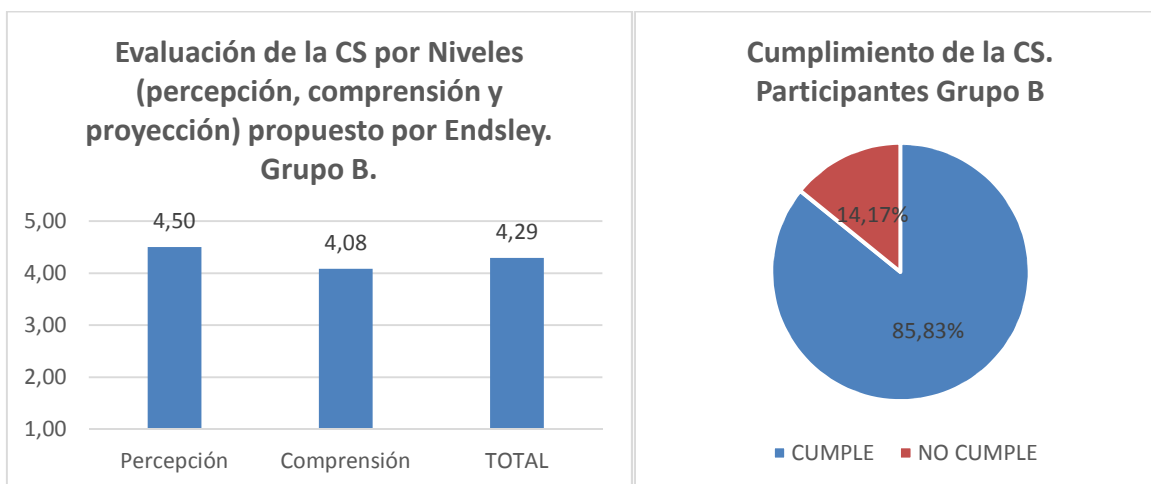


Figura 9.3 – Grafico de Evaluación de la CS por Niveles según Endsley: percepción, comprensión y proyección del Grupo B de participantes evaluados. Gráfico de cumplimiento de la CS de los participantes del Grupo B. (fuente: Propia)

Si analizamos de manera conjunta (Figura 9.4) los datos asociados a los dos grupos de población considerados (Grupo A y B) se observa que la valoración de la implementación evaluada es elevada (4,38) y que el 87% de los participantes consideró que el sistema

evaluado cubriría sus necesidades de información para percibir, comprender y, en su caso, tomar decisiones en el contexto recreado.

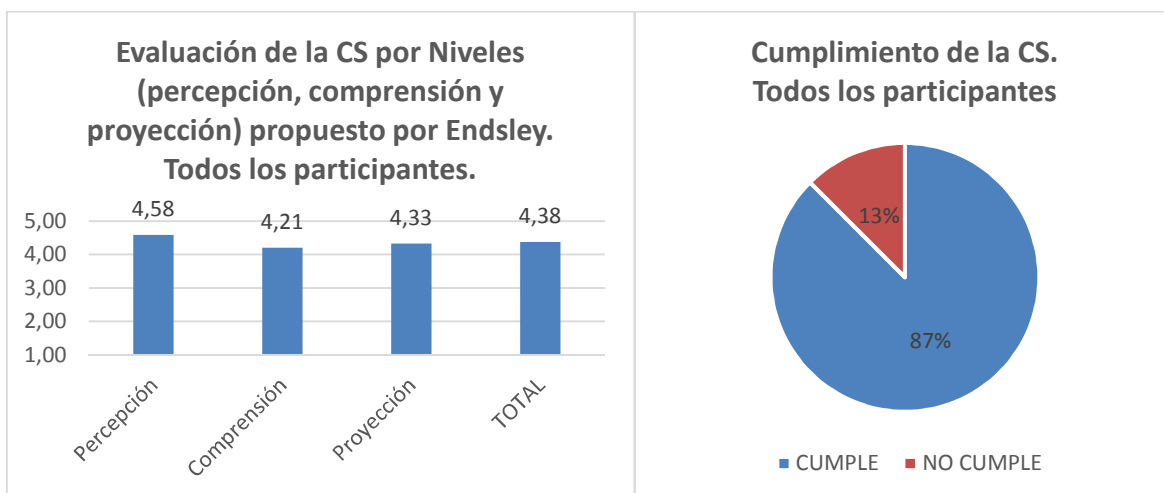
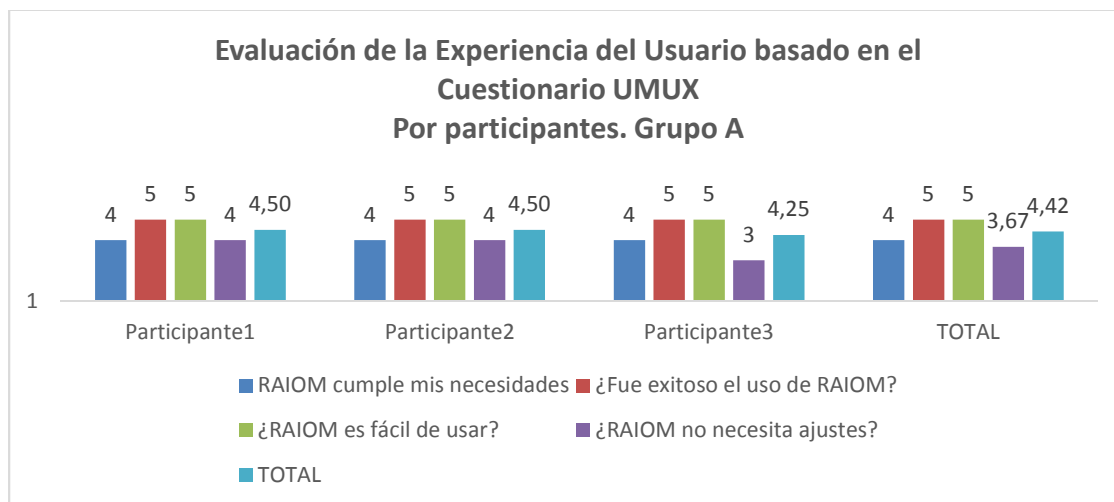


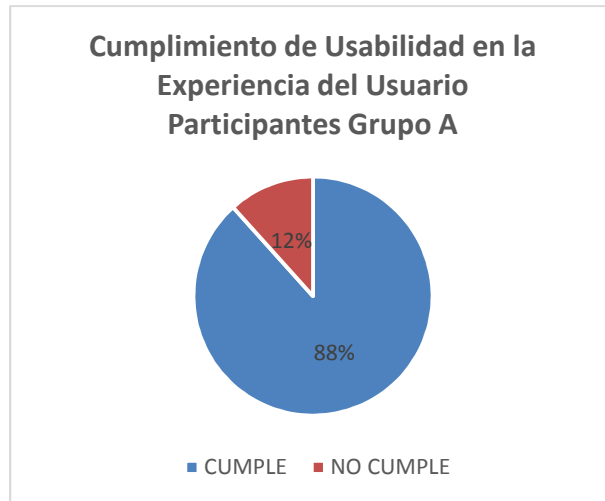
Figura 9.4 – Grafico de Evaluación de la CS por Niveles según Endsley: percepción, comprensión y proyección de todos los participantes evaluados (Grupo A y Grupo B). Gráfico de cumplimiento de la CS de todos los participantes (Grupo A y Grupo B). (fuente: Propia)

Finalmente, en la evaluación también se evaluó la experiencia del usuario (para el Grupo A ver Tabla 9.9 y la Figura 9.5 y para el Grupo B ver la Tabla 9.10 y la figura 9.6). Para ellos se pasó un cuestionario a los participantes basados en la propuesta UMUX. La valoración de los participantes del soporte que les proporciona RAIOM fue altamente positiva. El apartado peor valorado estuvo relacionado con los ajustes necesarios para el uso del sistema independientemente del grupo de participantes. Sin embargo, el sistema evaluado resultó agradable en su uso y fácil de usar una vez ya estaba ajustado.

Tabla 9.9 Cuadro de evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX. Grupo A

Participante Pregunta	P1	P2	P3	TOTAL
RAIOM cubre mis necesidades	4,00	4,00	4,00	4,00
RAIOM es agradable en uso	5,00	5,00	5,00	5,00
RAIOM es fácil de usar	5,00	5,00	5,00	5,00
RAIOM no necesita ajustes	4,00	4,00	3,00	3,67
TOTAL	4,50	4,50	4,25	4,42





**Figura 9.5 – Grafico de Evaluación de la Experiencia del Usuario basado en el cuestionario UMUX correspondiente a los participantes del Grupo A.
Gráfico de cumplimiento de Usabilidad en la Experiencia del Usuario correspondiente a los participantes del Grupo A.
(fuente: Propia)**

Tabla 9.10 Cuadro de evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX. Grupo B

Pregunta \ Participante	P1	P2	P3	TOTAL
RAIOM cubre mis necesidades	4,00	4,00	4,00	4,00
RAIOM es agradable en uso	4,00	5,00	5,00	4,67
RAIOM es fácil de usar	5,00	5,00	5,00	5,00
RAIOM no necesita ajustes	3,00	4,00	4,00	3,67
TOTAL	4,00	4,50	4,50	4,33

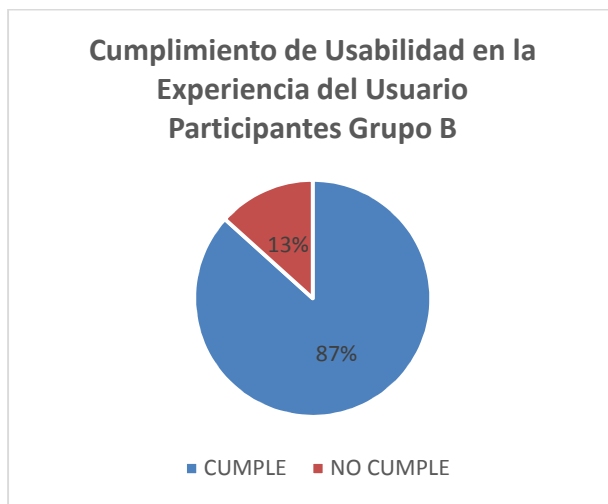
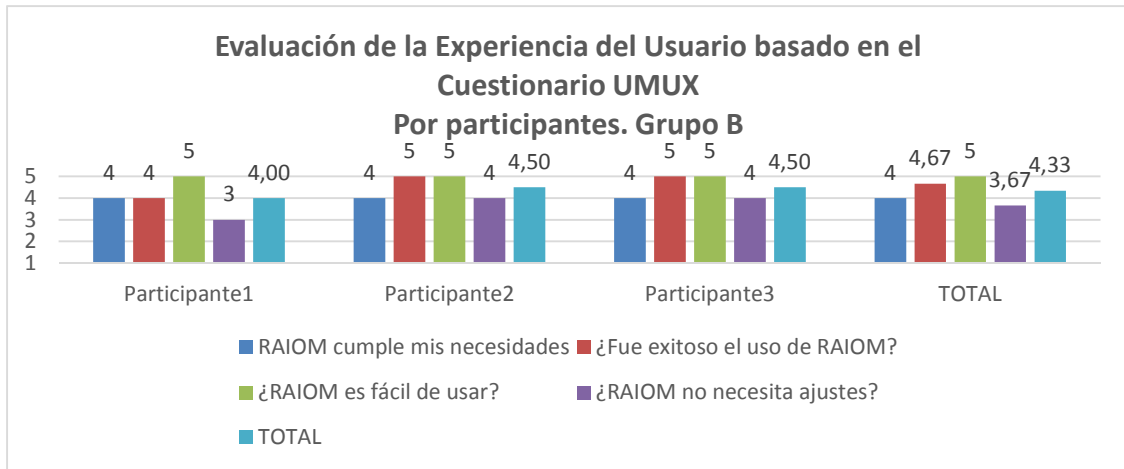
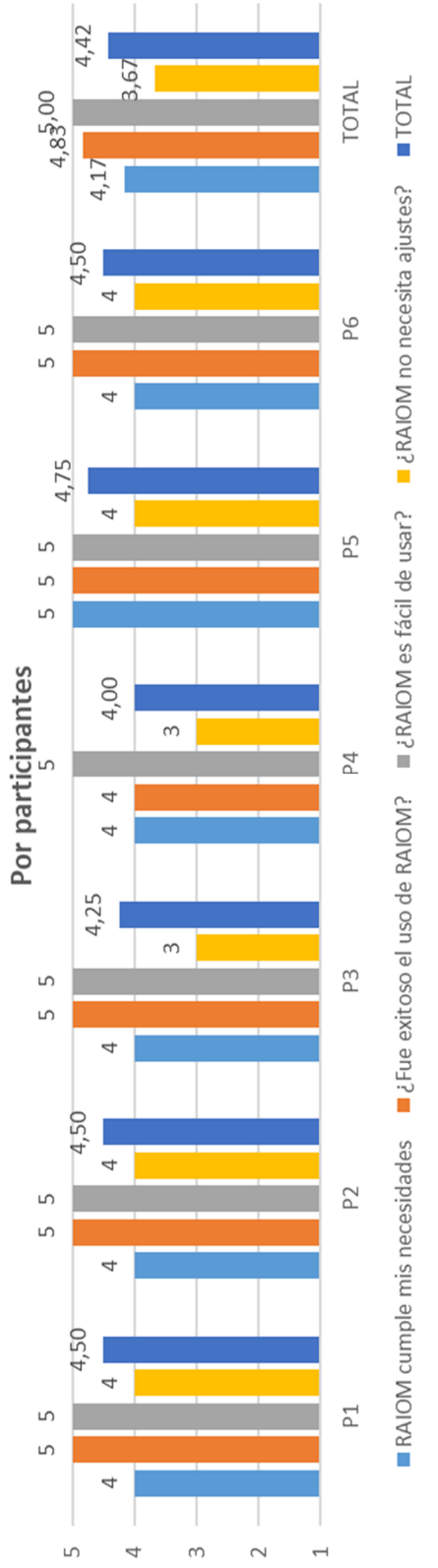


Figura 9.6 – Grafico de Evaluación de la Experiencia del Usuario basado en el cuestionario UMUX correspondiente a los participantes del Grupo B.
Gráfico de cumplimiento de Usabilidad en la Experiencia del Usuario correspondiente a los participantes del Grupo B.
(fuente: Propia)

Evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX



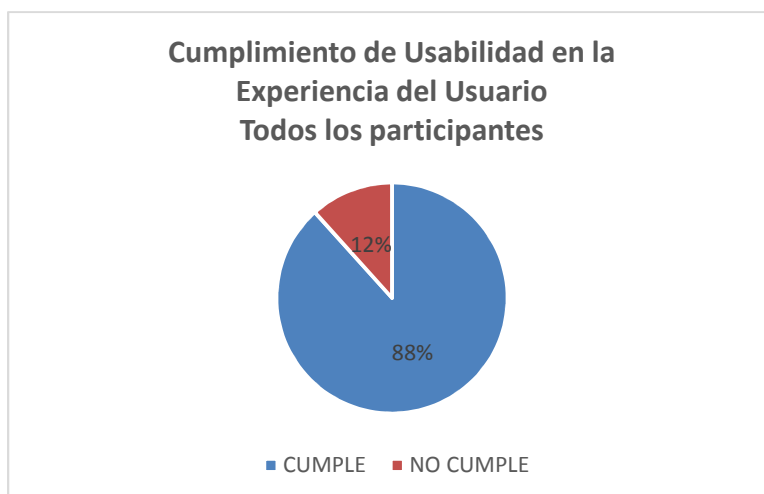


Figura 9.7 – Gráfico de Evaluación de la Experiencia del Usuario basado en el cuestionario UMUX correspondiente a todos los participantes (Grupo A y Grupo B). Gráfico de cumplimiento de Usabilidad en la Experiencia del Usuario correspondiente a todos los participantes (Grupo A y Grupo B). (fuente: Propia)

9.6.2 Prueba 2 realizada el 11 de noviembre de 2019

Se realizó un experimento adicional incorporando al mismo 4 participantes más, de los cuales un participante corresponde al Grupo A y los otros tres restantes al Grupo B. Cabe destacar que el protocolo que se utilizó para el experimento es el mismo que se detalla en la sección 9.2 - *Resumen de la evaluación*.

A continuación se detallan los resultados totales del primer experimento y del segundo experimento.

Tabla 9.11 Perfil de los 4 participantes adicionales al experimento inicial

Participante	Grupo	Establecimiento	Años de experiencia	Edad	Sexo	Nivel de uso de tecnología (alto/medio/bajo)
1	A	CITEDEF	>15	51	M	Medio
2	B	CITEDEF	>10	44	F	Medio
3	B	CITEDEF	>8	31	M	Alto
4	B	CITEDEF	>8	10	M	Alto

Tabla 9.12 Cuadro actualizado (P4) de Evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley. Grupo A.

Nivel	Pregunta	Puntaje por participante				
		P1	P2	P3	P4	Total
Percepción	Q1	5	5	5	5	
	Q2	5	4	5	4	
	Q3	5	5	5	5	
	Q4	5	5	2	4	
		5	4,75	4,25	4,50	4,63
Comprensión	Q5	5	4	4	4	
	Q6	4	4	5	4	
	Q7	5	5	5	5	
	Q8	3	4	4	3	
		4,25	4,25	4,5	4	4,25
Proyección	Q9	5	5	5	5	
	Q10	4	4	4	4	
	Q11	4	3	5	5	
		4,33	4	4,67	4,67	4,42

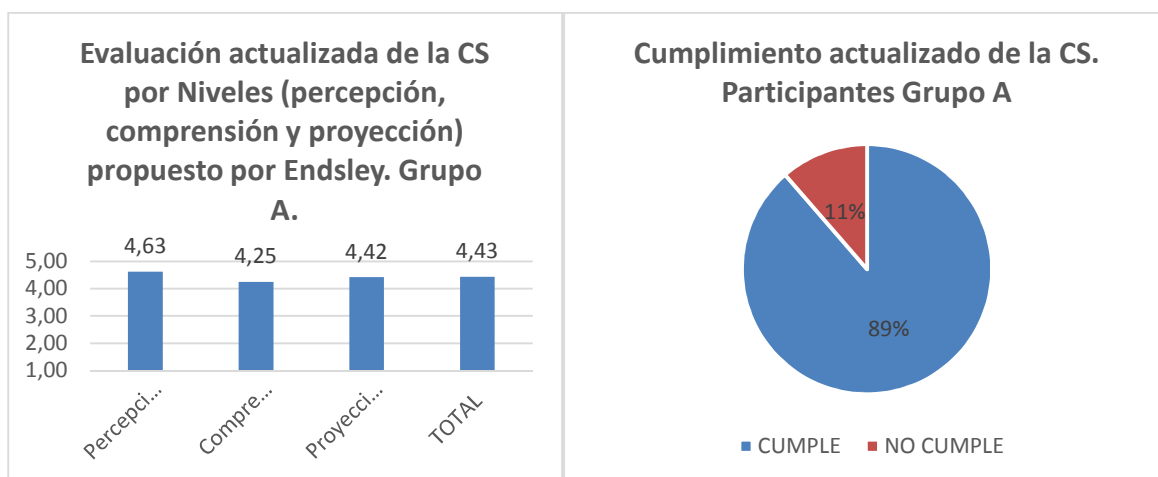


Figura 9.8 – Grafico de Evaluación actualizado de la CS por Niveles según Endsley: percepción, comprensión y proyección del Grupo A de participantes evaluados. Gráfico de cumplimiento actualizado de la CS de los participantes del Grupo A. (fuente: Propia)

Tabla 9.13 Cuadro actualizado (P4, P5 y P6) de Evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley. Grupo B.

Nivel	Pregunta	Puntaje por participante						
		P1	P2	P3	P4	P5	P6	Total
Percepción	Q1	4	4	5	5	5	5	
	Q2	5	5	4	5	5	5	
	Q3	5	5	5	5	5	4	
	Q4	4	5	3	5	5	4	
		4,5	4,75	4,25	5	5	4,50	4,67
Comprensión	Q5	4	3	5	3	3	3	
	Q6	4	4	5	5	4	4	
	Q7	5	4	5	5	4	4	
	Q8	4	2	4	3	2	3	
		4,25	3,25	4,75	4	3,25	3,50	3,83

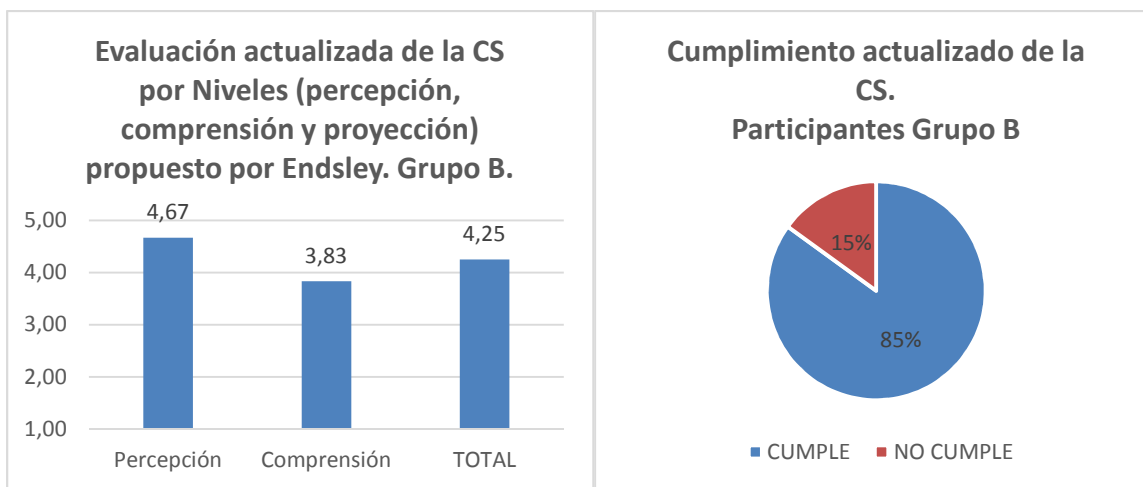


Figura 9.9 – Grafico de Evaluación actualizado de la CS por Niveles según Endsley: percepción, comprensión y proyección del Grupo B de participantes evaluados. Gráfico de cumplimiento actualizado de la CS de los participantes del Grupo B. (fuente: Propia)

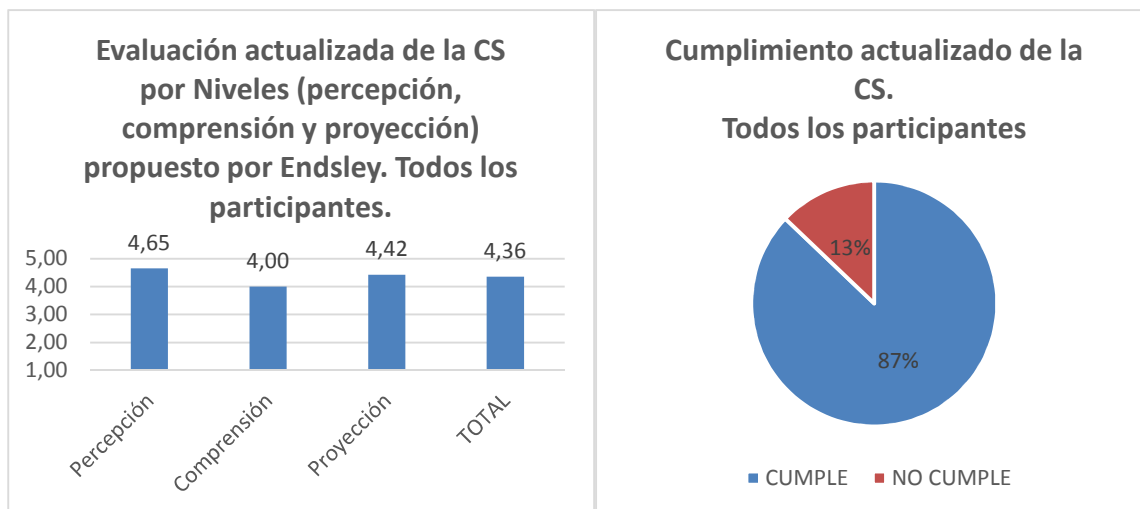


Figura 9.10 – Grafico de Evaluación actualizado de la CS por Niveles según Endsley: percepción, comprensión y proyección de todos los participantes evaluados (Grupo A y Grupo B).

**Gráfico de cumplimiento actualizado de la CS de todos los participantes (Grupo A y Grupo B).
(fuente: Propia)**

Por otro lado también se actualizó la evaluación de la experiencia del usuario utilizando el protocolo UMUX:

Tabla 9.14 Cuadro actualizado (P4) de evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX. Grupo A

Pregunta \ Participante	P1	P2	P3	P4	TOTAL
RAIOM cubre mis necesidades	4	4	4	4	4,00
RAIOM es agradable en uso	5	5	5	4	4,75
RAIOM es fácil de usar	5	5	5	5	5,00
RAIOM no necesita ajustes	4	4	3	3	3,50
TOTAL	4,50	4,50	4,25	4,00	4,31

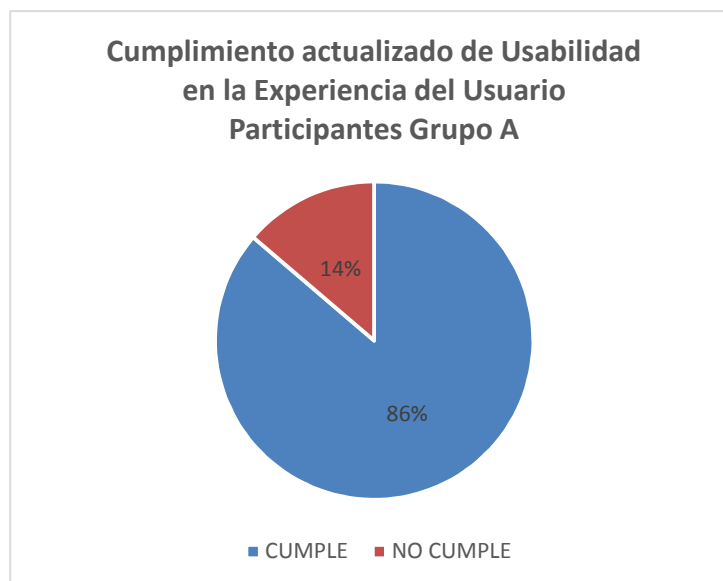
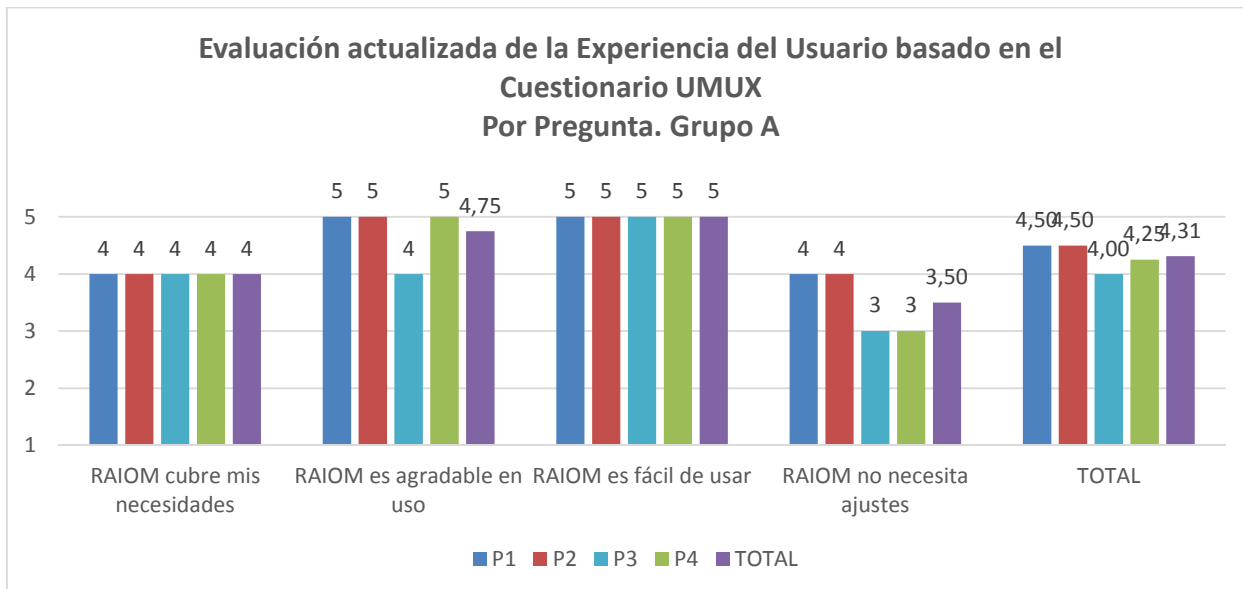
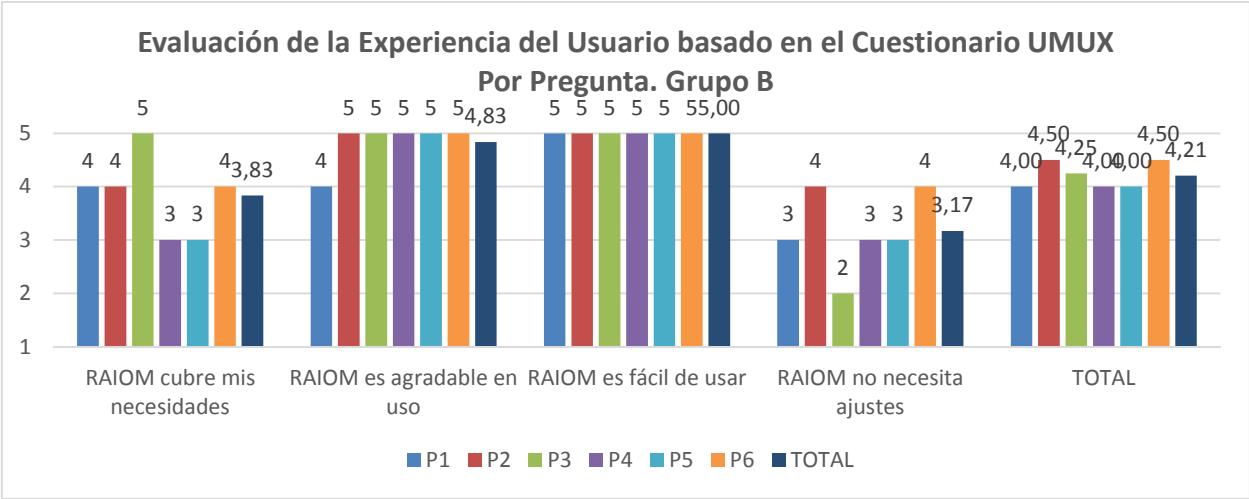
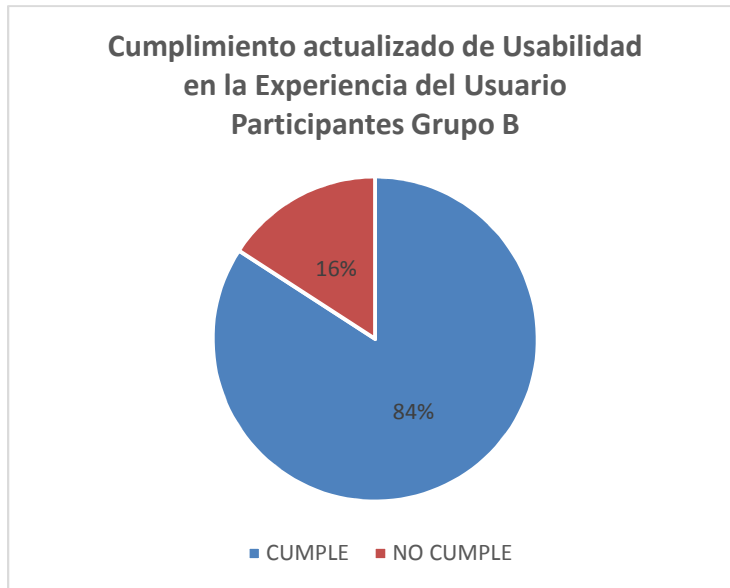


Figura 9.11 – Grafico actualizado de Evaluación de la Experiencia del Usuario basado en el cuestionario UMUX correspondiente a los participantes del Grupo A. Gráfico actualizado de cumplimiento de Usabilidad en la Experiencia del Usuario correspondiente a los participantes del Grupo A. (fuente: Propia)

Tabla 9.15 Cuadro actualizado (P4, P5, P6) de evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX. Grupo B

Pregunta \ Participante	P1	P2	P3	P4	P5	P6	TOTAL
RAIOM cubre mis necesidades	4	4	5	4	3	3	3,83
RAIOM es agradable en uso	4	5	5	5	5	5	4,83
RAIOM es fácil de usar	5	5	5	5	5	5	5,00
RAIOM no necesita ajustes	3	4	2	4	3	3	3,17
TOTAL	4,00	4,50	4,25	4,50	4,00	4,00	4,21





**Figura 9.12 – Grafico actualizado de Evaluación de la Experiencia del Usuario basado en el cuestionario UMUX correspondiente a los participantes del Grupo B.
Gráfico actualizado de cumplimiento de Usabilidad en la Experiencia del Usuario correspondiente a los participantes del Grupo B.
(fuente: Propia)**

A continuación se detallan los totales globales del experimento actualizado (prueba 1 y prueba 2) con sus correspondientes resultados:

Tabla 9.16 Resultados globales actualizados de la evaluación de la CS por Niveles (percepción, comprensión y proyección) propuesto por Endsley

Nivel Grupo	Percepción	Comprensión	Proyección	Total
A	4,63	4,25	4,42	4,43
B	4,67	3,83	N/A	4,25
Total	4,65	4,00	4,42	4,36



Figura 9.13 – Gráfico de cumplimiento actualizado de la CS de todos los participantes (Grupo A y Grupo B). (fuente: Propia)

Tabla 9.17 Resultados globales actualizados de la evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX

Grupo	Total
A	4,31
B	4,21
Total	4,26

Actualizando los resultados en comparación con la primera evaluación aún podemos mencionar que con estos resultados se puede ser razonablemente optimista con el grado de consecución de la CS obtenido por los participantes en el experimento, utilizando el equipo y realizando las tareas propuestas con efectividad y eficiencia.

Los participantes que pertenecen al Grupo A, el conjunto de individuos decisores, representa un cumplimiento de la CS del 89%, mientras que para los usuarios del Grupo B representa un cumplimiento del 85%. Al igual que la primera evaluación, las facilidades de percepción son superiores a las logradas para otros niveles de Endsley (comprensión y proyección). Pero esos otros niveles de Endsley también se responden con un grado alto de efectividad.

En cuanto a la experiencia del usuario (UX) se llega a la conclusión que la interfaz diseñada cumple en un 86% para los usuarios pertenecientes al Grupo A y un 84% para los usuarios del Grupo B. Por lo tanto, se ratifica que los participantes se sienten satisfechos con la implementación del sistema RAIOM facilitada en el experimento y que el mismo les parece útil.

9.7 Resumen del capítulo

Este capítulo presentó la evaluación del framework propuesto a través del desarrollo de una aplicación donde se integraban todos los recursos tecnológicos mencionados en capítulos anteriores. Para llevar a cabo la evaluación de efectividad, eficacia y usabilidad del artefacto construido, se dio cuenta del uso de técnicas tales como SAGAT y SART complementadas mediante el despliegue de User Testing y Thinking Aloud protocol y, adicionalmente, se medió la experiencia del usuario (UX) en la utilización del despliegue probado. Las pruebas fueron organizadas de acuerdo al protocolo detallado en la ISO/IEC 25062 (2006). Un resumen de los resultados obtenidos de las evaluaciones realizadas se encuentra en la sección *9.2 - Resumen de la evaluación* del presente capítulo.

Capítulo 10

Conclusiones y trabajos futuros

10.1 Conclusiones

En este capítulo se detallan las principales conclusiones alcanzadas y los workflows / lineamientos realizados durante la elaboración de esta tesis doctoral. En este sentido, en el Capítulo 1, en la sección 1.4 - *Metodología de investigación* se identificó la metodología DSR como elemento conductor de este trabajo. Por ello ahora se repasan seguidamente, y de acuerdo con dicha metodología, los principales workflows abordados:

1) Diseñar un artefacto

El objetivo principal del presente trabajo de tesis es desarrollar e implementar un framework propio para ayudar y apoyar a los militares a mejorar su CS en ambientes hostiles y desconocidos. Una vez que se llegó a la conclusión que era necesario desarrollar el framework, se dio comienzo al diseño del mismo considerando una arquitectura de capas y componentes a fin de facilitar la integración tecnológica.

Se hizo hincapié en implementar algunas técnicas de optimización con el objeto de ejecutar procesos asincrónicos.

Paralelamente al diseño del framework se analizó el esquema conceptual de operaciones militares, se identificaron sus actores y sus funciones. Se centró en identificar y analizar los requisitos funcionales a fin de realizar el diseño de las aplicaciones de concepto y sus interfaces. Un aporte importante fue la el diseño de un modelo tridimensional que sirvió para la identificación, selección y clasificación de requisitos de la CS. Para ello se combinaron el modelo de 3 niveles propuesto por Endsley (percepción, comprensión y proyección), el modelo de Gutwin et al. [70], y el *ámbito* de aplicación. De esta manera se construyó un modelo de 3 ejes en el espacio. Cada eje está formado por los aspectos de cada modelo antes mencionado más el eje del *ámbito* de

aplicación, es decir, *Visual, Auditivo y Háptico* que corresponden a los 3 modos de interacción con el sistema.

Se evaluó qué metodología era la apropiada para abordar el diseño del framework y el diseño de las interfaces de las aplicaciones de prueba de concepto llegando a la conclusión que SCRUM+DCU era la adecuada. El objetivo era utilizar un método ágil (SCRUM) y centrada en el usuario (DCU), es decir, se analizó la utilización de una metodología en donde el usuario sea el centro del diseño de las interfaces y de la interpretación y uso de las funcionalidades principales que debería soportar el framework.

Posteriormente se desarrolla una aplicación utilizando el framework propuesto a fin evaluar las características técnicas, funcionales del artefacto y la eficacia y eficiencia de los despliegues desarrollados.

El framework desarrollado tiene las siguientes características centrales:

- Arquitectura cliente-servidor
- Procesamiento distribuido entre dispositivo de visualización (gafas de RA, smartphone, tableta, hand-held) y micro computadora ODROID-XU3
- Visualización de alertas (iconografía y simbología militar) de amenazas en un mapa de 360 grados
- Reconocimiento de construcciones edilicias, previamente almacenadas en una base de datos
- Reconocimiento gestual y vocal para la entrada de datos (mandos de acciones predefinidas)
- Tracking del recorrido realizado

El artefacto y su desarrollo se encuentran detallados en la *parte II - Framework Propuesto*.

2) Relevancia del problema (dominio del problema)

Se puso foco en entender qué es la CS y se analizó un conjunto de proyectos militares que utilizan la RA como soporte tecnológico para mejorar la CS. De este análisis se llegó a la conclusión que es insuficiente para cumplir con los requisitos específicos y la hipótesis de esta tesis.

Para lograr los objetivos establecidos se realizó una revisión sistemática que constó en evaluar e interpretar las investigaciones relevantes disponibles utilizando una metodología

confiable, rigurosa y auditable. Para ello se utilizó la metodología propuesta por [86]. Adicionalmente se realizaron tablas comparativas entre los dispositivos digitales que utilizan la RA como método de visualización para soportar a la CS en situaciones hostiles para determinar la necesidad de construir un artefacto propio (ver sección 4.4 - *Comparativa de tecnologías digitales militares que utilizan RA*). Adicionalmente, se realizó una comparación de los frameworks de uso civil que reunían características aproximadas a las necesidades planteadas (ver sección 5.3 - *Comparativa de frameworks de RA más utilizados*). Una revisión de la tecnología aplicada a proyectos militares como también la existencia de frameworks no militares, arquitecturas implementadas para el diseño de las mismas, encuestas realizadas a personal militar, entre las principales tareas se puede encontrar en la *parte I - Estado de Situación*.

3) Evaluación del Diseño

Se evaluó un conjunto de características del framework propuesto en los diferentes despliegues que se han implementado. El objetivo de la evaluación fue determinar que la implementación realizada asegura niveles altos de CS para percibir, comprender y tomar decisiones adecuadas con efectividad y eficiencia y con alto grado de experiencia de usuario (UX). Para llevar a cabo las actividades de evaluación fueron utilizadas las técnicas SAGAT y SART (*Capítulo 9*), complementada con un cuestionario UMUX para evaluar la experiencia del usuario. La documentación de las actividades de evaluación se realizó teniendo en cuenta la norma ISO/IEC 25062 (2006).

Dichas características estuvieron relacionadas con evaluar la CS utilizando los 3 niveles propuestos por Endsley: percepción, comprensión y proyección y se analizó la experiencia del usuario en utilizar la interfaz desarrollada. Las actividades que ejecutaron los participantes fueron seleccionadas teniendo como soporte 2 premisas, por un lado, los 3 niveles propuestos por Endsley y, por el otro lado, se utilizó la taxonomía de Bloom que identifica los procesos cognitivos haciendo hincapié en los de orden inferior, es decir, aquellos relacionados con recordar, comprender y aplicar.

Con los resultados obtenidos se puede observar que el grado de consecución de la CS obtenido fue eficiente y eficaz en las tareas ejecutadas por los participantes. Se destacó que las facilidades de percepción son superiores a las logradas para otros niveles de Endsley (comprensión y proyección). No obstante esos otros niveles de Endsley también se responden con un grado alto de efectividad.

En cuanto a la experiencia del usuario (UX) se puede manifestar que los participantes se sienten satisfechos con la implementación del sistema RAIOM facilitada en el experimento y que el mismo les parece útil.

4) Contribución de la investigación

Las contribuciones realizadas en el ámbito de la CS resultan reveladoras tanto en el plano teórico donde hemos presentado una caracterización o modelo que permite organizar las soluciones de diseño asociadas a diferentes estilos de interacción que persiguen el mismo propósito: potenciar la CS. En el plano tecnológico se ha propuesto y demostrado la validez de una arquitectura multicapa que posibilita el desarrollo actual y futuro de soluciones, basadas en la RA, para abordar los retos actuales que propone la implementación del Soldado del Futuro.

Las contribuciones aportadas en el presente trabajo de tesis se detallan en la sección *10.2 Aportes realizados*

5) Rigor de la investigación

La metodología utilizada en esta etapa para diseñar el artefacto y realizar, posteriormente, su evaluación se menciona en la sección *1.4 - Metodología de investigación* y se detalla en la presente sección. Se describe la arquitectura de software utilizada, el diagrama de clases y la arquitectura del sistema, la integración entre los dispositivos y la integración de las minicomputadoras ODROID en una arquitectura distribuida. Por otro lado el método de evaluación del artefacto fue acordado con personal militar y se encuentra detallado en la sección *Capítulo 9*.

6) Diseño como un proceso de búsqueda

Se debe tener en cuenta la naturaleza iterativa del proceso de diseño (SCRUM+DCU). Las pruebas de concepto denotan el proceso evolutivo del artefacto, en cada etapa de prueba se detallan las funcionalidades relevantes implementadas en los despliegues correspondientes. La iteración final se encuentra representada por la integración de todas las pruebas en una aplicación final, sección *8.6 - Integración de tecnologías*.

7) Comunicación de la investigación

Parte de las publicaciones realizadas están orientadas al ámbito militar teniendo como base la CS como problemática a resolver utilizando el framework que se construyó a partir de este trabajo de tesis. No obstante, también se han publicado artículos académicos centrándose

principalmente en los proyectos de investigación llevados a cabo por doctorandos y estudiantes de las especializaciones vinculadas al estudio de procesamiento de imágenes, visión por computador, realidad aumentada y realidad virtual realizadas en diferentes Universidades nacionales. Dichas publicaciones se encuentran detalladas en la sección *10.3 - Publicaciones*.

Dado que la premisa de esta investigación fue impulsada por la necesidad de resolver un problema utilizando un artefacto tangible, el framework propuesto y su evaluación proporciona las herramientas conceptuales para dar forma a la investigación, así como una base sólida para discutir las diferentes partes del proyecto de investigación.

Como conclusión personal debo destacar en principio la colaboración recibida por parte de los integrantes del grupo de trabajo del que pertenezco (GEINF - CITEDEF) y del personal militar (EE.AA) que fue clave para definir los requerimientos y evaluar los despliegues desarrollados. La presente tesis dio origen al proyecto RAIOM cuyo principal usuario es el EE.AA y los grupos comandos que lo utilizan. Cabe destacar que el desarrollo de la aplicación es considerado único en Latinoamérica y me enorgullece haber sido el director técnico del proyecto y haber capacitado a un equipo de investigadores que hoy forman parte del departamento de I+D+i de Tecnología Informática del CITEDEF y son referentes a nivel nacional en procesamiento de imágenes y visión por computador utilizando a la RA como tecnología principal en el desarrollo del presente trabajo de investigación. El proyecto de investigación dio origen a un conjunto de artículos y notas técnicas (CITEDEF) que fueron presentadas entre los años 2013 y 2018 dando a conocer los avances que se iban realizando en el proyecto. Adicionalmente, y como consecuencia de la investigación y realización de la presente tesis, se presentó un proyecto de I+D para diseñar un framework multipropósito de Realidad Aumentada (RA) y Visión Artificial (VA) bajo mi dirección. Dicho proyecto, actualmente, se está llevando a cabo en el Laboratorio de Procesamiento de Imágenes y Visión Artificial de la Facultad de Ingeniería y Tecnología Informática de la Universidad de Belgrano (FITI – UB). El framework servirá como base para la construcción de aplicaciones móviles y de escritorio para que puedan ser utilizadas en diferentes disciplinas tales como medicina, entretenimiento, industria, turismo y educación, entre las principales. Uno de los objetivos centrales del proyecto de investigación es la incorporación de conocimientos sólidos en el uso de tecnologías de RA y VA aplicable a cualquier ámbito de la industria y de la vida social. Otro de los objetivos principales es formar a alumnos en las tareas de investigación y desarrollo de proyectos informáticos de avanzada. En lo que concierne a la educación este tipo de tecnología

pondrá foco principalmente en reforzar el aprendizaje a través del contenido y a través de imágenes temáticas, entre otros, logrando la inclusión y el entendimiento de personas de diferentes extractos sociales. Por otro lado, y en el mismo claustro, se creó un curso de posgrado de Educación Continua denominado “Interfaces Avanzadas y Visión por Computador: de lo real a lo digital”, cuyo objetivo se centra en capacitar en tecnología aplicada al procesamiento de imágenes, visión por computador e interfaces avanzadas.

10.2 Aportes realizados

En cuanto al aporte de la presente tesis se pueden listar las principales contribuciones:

- a) El uso de la RA como soporte a la CS para la toma de decisiones en el sector militar aprovechando el avance del software y de los dispositivos móviles actuales;
- b) La utilización de una metodología combinada SCRUM-DCU para ello el diseño del framework y el desarrollo de las aplicaciones;
- c) El diseño de un modelo tridimensional (3D-SAM) para la identificación, selección y clasificación de requisitos de la CS y diseño de prototipos;
- d) La consideración de una arquitectura basada en capas a fin facilitar la integración tecnológica;
- e) La utilización de métodos válidos y confiables para evaluar la CS tales como SAGAT, SART, User Testing y Thinking Aloud y la experiencia del usuario mediante UMUX;
- f) La implementación de un sistema distribuido utilizando componentes de hardware externo (ODROID-XU3) en donde se ejecuta el procesamiento de imágenes y reconocimiento de objetos.

Adicionalmente el trabajo de tesis sigue lo planteado en [14], es decir, de los cuatro niveles de tareas para el procesamiento de RA entre el cliente y el servidor, se aplica el nivel tres, es decir, en el servidor se ejecutan las tareas de tracking y de procesamiento de imágenes mientras que en el cliente se ejecutan las tareas de captura de video, rendering y visualización (display).

Como se mencionó anteriormente, los aportes llevan a preguntar cuáles son las diferencias existentes con otros frameworks ya sea para uso civil o militar. Por tal motivo la pregunta que se formula de este análisis es la siguiente:

¿Qué diferencia existe entre los componentes digitales militares que utilizan la RA como método de visualización y el diseño y desarrollo de un framework específico propio de RA que le

sirva a las FF.AA desarrollar sus propias aplicaciones para detectar amenazas y mejorar la CS del operador? (sección 4.4 - Comparativa de tecnologías digitales militares que utilizan RA)

- Especifico y centrado para tareas de inteligencia, vigilancia y reconocimiento (ISR colectivamente) militar (inteligencia militar)
 - o Detección y reconocimiento de objetos
- Centrado en mejorar la CS del operador
 - o Simbología militar de amenazas circundantes
- Altamente interactivo
 - o Reconocimiento gestual y vocal para la toma de datos
- Procesamiento distribuido
 - o Implementado en una arquitectura cliente / servidor. El computo de los algoritmos de reconocimiento de objetos se realizan en el servidor
- Implementado con componentes del tipo COTS¹⁵ de bajo costo de adquisición
- Fuerte integración con sensores externos para la toma de datos del entorno

Se puede agregar que tienen una fuerte orientación a:

- Conocimiento de las amenazas del contexto
- Soportar la CS del operador

Cabe aclarar que el trabajo de tesis se realizó en el laboratorio del LIDI de la UNLP en conjunto con el laboratorio de investigación del CITEDEF, con el apoyo de las EE.AA para la concreción del presente trabajo de tesis doctoral.

10.3 Publicaciones

La presente tesis doctoral está avalada por las siguientes publicaciones:

¹⁵ COTS: Commercial Off-The-Shelf: es un término del Reglamento Federal de Adquisiciones (FAR), que define un elemento no-desarrollativo (NDI) de suministro, que es a la vez comercial y se vende en grandes cantidades en el mercado comercial, y que puede ser adquirido o utilizado bajo contrato gubernamental de la misma forma exacta a como está disponible al público en general.

Revistas Nacionales e Internacionales

- Autores: Alejandro Mitaritonna, María José Abásolo, Francisco Montero Simarro
Título: An Augmented Reality-based framework to support Military Situational Awareness
Revista: Defense Science Journal (DSJ). ISSN 0011-748X
Estado: Enviado el 09/08/2019 (en proceso de evaluación)
- Autores: Alejandro Mitaritonna
Título: Tecnologías emergentes en la educación: la realidad aumenta
Revista: Perspectivas: Revista Científica de la Universidad de Belgrano, V 1, N^a 2, 2018. Págs. 85-93.
Estado: Publicado 15/11/2018

Congresos Internacionales:

- Autores: Alejandro Mitaritonna, María José Abásolo, Francisco Montero Simarro
Titulo: Situational Awareness through Augmented Reality: 3D-SA Model to relate Requirements, Design and Evaluation
Congreso: The 9th International Conference on Virtual Reality and Visualization (ICVRV 2019). SHENZHEN UNIVERSITY TOWN CONFERENCE CENTER, Shenzhen, China
Estado: Aceptado. IEEE Proceedings of The 9th International Conference on Virtual Reality and Visualization (ICVRV 2019).
- Autores: Mitaritonna Alejandro; Abásolo María José
Título: Improving Situational Awareness in Military Operations using Augmented Reality
Congreso: 23rd International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision 2015, Pilsen, República Checa. Poster
ISBN 978-80-86943-67-1
- Autores: Mitaritonna, Alejandro; Abásolo, María José;
Título del capítulo: Mejorando la conciencia situacional en operaciones militares utilizando la realidad aumentada.
Libro de actas: Fundación de altos estudios en ciencias exactas. 2013. P356 - 365.
Congreso: CACIC 2013 XIX Congreso Argentino de Ciencias de la Computación.
ISBN 978-987-23963-1-2

Congresos Nacionales:

- Autores: María José Abásolo, Armando De Giusti, Marcelo Naiouf, Patricia Pesado, Cecilia Sanz, Sebastián Barbieri, Ramiro Boza, Wilma Gavilanes, Alejandro Mitaritonna, Nahuel Prinscich, Mario Alberto Vincenzi, Francisco Montero, Francisco Perales López
Título: Aplicaciones de Realidad Virtual, Realidad Aumentada e Interfaces Multimodales
Congreso: Workshop de Investigadores en Ciencias de la Computación (WICC 2019); Universidad Nacional de San Juan, San Juan, 2019
Estado: Enviado el 13/03/2019 (en proceso de evaluación)
- Autores: Alejandro Mitaritonna, Juan Lestani, Silvana Olmedo, Leonardo Inza
Título: Investigación y desarrollo de aplicaciones de Realidad Aumentada y Visión por Computador utilizando un framework multipropósito
Congreso: Workshop de Investigadores en Ciencias de la Computación (WICC 2019); Universidad Nacional de San Juan, San Juan, 2019
Estado: Enviado el 13/03/2019 (en proceso de evaluación)
- Autores: María José Abásolo, Alejandro Mitaritonna, Sebastián Castañeda, Cecilia Sanz, Ramiro Boza, Nahuel Prinscich, Telmo Silva, Magdalena Rosado, Marcelo Naiouf, Patricia Pesado, Armando De Giusti
Título: Aplicaciones de Visión por Computador, Realidad Aumentada y TVDi
Congreso: Workshop de Investigadores en Ciencias de la Computación (WICC 2018); Ciudad de Corrientes, 2018
ISBN 978-987-3619-27-4
- Autores: Alejandro Mitaritonna, Juan Lestani, Francisco Tarulla, Tomás Poeta, Silvana Olmedo, Carolina Páez, Martín Lorenzo, Florencia Vela, Leonardo Inza
Título: Realidad Aumentada y Visión por Computador. Framework multipropósito
Congreso: Workshop de Investigadores en Ciencias de la Computación (WICC 2018); Ciudad de Corrientes, 2018
ISBN 978-987-3619-27-4
- Autores: Alejandro Mitaritonna, Juan Lestani
Título: Framework multipropósito de Realidad Aumentada y de Visión Artificial

Congreso: XII Congreso de Tecnología en Educación y Educación en Tecnología (TE&ET 2017); Universidad Nacional de La Matanza, 2017

ISBN 978-987-4417-04-6

- Autores: María José Abásolo, Alejandro Mitaritonna, María José Bouciguez, Natalia Encina, Mario Vicenzi, Armando De Giusti, Marcelo Naiouf, Javier Giacomantone, Cristina Manresa Yee

Título: Realidad Aumentada, Realidad Virtual, Interfaces Avanzadas y Juegos Educativos

Congreso: XVIII Workshop de Investigadores en Ciencia de la Computación (WICC 2016); Concordia, Entre Ríos, 2016

ISBN 978-950-698-377-2

- Autores: Abásolo María José; Mitaritonna Alejandro; Encina Natalia; Vicenzi Mario; Borelli Lucas; De Giusti Armando; Naiouf Marcelo; Giacomantone Javier

Título: Realidad Aumentada y Realidad Virtual

Congreso: XVII Workshop de Investigadores en Ciencia de la Computación (WICC 2015); Salta, 2015

- Autores: María José Abásolo, Alejandro Mitaritonna, Javier Giacomantone, Armando de Giusti, Marcelo Naiouf, Francisco Perales, Cristina Manresa, Marcelo Vénere, Cristian García Bauza

Título: Realidad Virtual, Realidad Aumentada y TVDI

Congreso: WIIC 2014 XVI Workshop De Investigadores en Ciencias de la Computación.

Instituto de Desarrollo Económico e Innovación de la Universidad Nacional de Tierra del Fuego, Antártida e Islas Del Atlántico Sur

- Autores: Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves

Título: RAIOM. Introducción a la arquitectura del framework de Realidad Aumentada

Congreso: VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014) e140109

- Autores: Alejandro Mitaritonna, Lucas Pandolfo, Dario Yokhdar y Carlos Esteves

Título: RAIOM. Introducción a los algoritmos de visión por computador

Congreso: VI Congreso y Workshop Argentino en Ciencias de las Imágenes (ECIMAG 2014) e140110

- Autores: Abásolo, María José; Mitaritonna, Alejandro; Giacomantone, Javier; de Giusti, Armando; Naiouf, Marcelo; Manresa, Cristina; Perales, Francisco; Sansó Más, Ramón; Castro, Silvia Mabel; Caluva, Claudio; Aranguren, Silvia; Muzachiodi, Rodolfo.

Título: visión por computador e informática gráfica. Realidad virtual, realidad aumentada e interfaces avanzadas.

Congreso: WIIC 2013 XV Workshop de Investigadores en Ciencias de la Computación.

Universidad Autónoma de Entre Ríos (UADER). 2013. P274 - 279.

ISBN 978-987-281-796-1

Informes Técnicos:

- Mitaritonna, Alejandro; Esteves, Carlos Alberto; Yokhdar, Dario Oscar.

Utilización De Mapas Off-Line Para Geolocalización Utilizando Fuentes De Red Externa.

Nota Técnica. Español. Argentina. 2014. Citedef (Geinf).

Informe técnico: Nro: 922

- Mitaritonna, Alejandro; Esteves, Carlos Alberto; Yokhdar, Dario Oscar.

Diseño De Un Modelo De Representación Visual Para Operaciones Militares.

Nota Técnica. Español. Argentina. 2014. Citedef (Geinf).

Informe técnico Nro: 923

- Mitaritonna, Alejandro; Esteves, Carlos Alberto; Yokhdar, Dario Oscar.

Identificación De Aviones De Combate Utilizando Técnicas De Procesamiento De Imágenes En Un Ambiente Controlado.

Nota Técnica. Español. Argentina. 2014. Citedef (Geinf).

Informe técnico Nro: 927

Otras Publicaciones:

- Autores: Alejandro Mitaritonna, María José Abásolo

Título: Hardware and software considerations for RAIOM Mobile Augmented Reality Framework

Publicado: Academia.edu

Acceso:http://www.academia.edu/7948407/Hardware_and_software_considerations_for_RAIOM_Mobile_Augmented_Reality_Framework

Año: 2013

10.4 Trabajos futuros

Como línea de trabajo futura, se propone la posibilidad de desarrollar un conjunto de funcionalidades que no fueron implementadas en el framework propuesto ya sea por tiempos y/o por el alcance definido. Estas funcionalidades están relacionadas con potenciar el uso de la CS y brindar un soporte más amplios y efectivo en los procesos de toma de decisiones.

Estas funcionalidades adicionales permitirán ampliar el uso pues las aplicaciones que extiendan del framework estarán preparadas para ser utilizadas en misiones en donde, por ejemplo, la falta de luz visible es un factor clave en las operaciones nocturnas.

Por otro lado, la utilización de mecanismos más efectivos de clasificación conlleva a pensar en nuevas herramientas que sean incorporadas en la estructura middleware del framework.

Más allá del framework como producto software, integrar en un HMD militarizado cada componente de hardware propuesto en el presente informe.

Desde el punto del experimento y de la evaluación cabe mencionar que tiene ciertas limitaciones por lo que se aconseja ampliar el número de usuarios que participen en las evaluaciones, entre otras.

Por lo tanto, y de acuerdo a lo expresado anteriormente, existen diversas líneas de investigación que quedan abiertas luego de la finalización del informe, a saber:

- *Capacidad de reconocimiento de objetos en modalidad nocturno*: esta funcionalidad implicaría desarrollar un Módulo para operaciones nocturna implementando algoritmos específicos en el módulo de *Vision*. De esta manera se ampliará la capacidad del módulo lo que posibilitará desarrollar aplicaciones extendiendo del framework esta característica. Se podrán desarrollar aplicaciones que requieran la utilización de reconocimiento de objetos bajo condiciones de poca luz o la carencia total de ésta.
- *Deep Learning (DL)*: se propone la implementación de técnicas de clasificación basado en la utilización de redes neuronales implementada en el módulo de *Vision* o en su defecto en un módulo específico de DL diseñado separadamente. La idea principal es sustituir los algoritmos de clasificación utilizados en el framework actual por técnicas de DL tales como Convolutional Neural Network (CNN) para el procesamiento de imágenes. Se propone CNN ya que debido a que su aplicación es realizada en matrices bidimensionales, son muy

efectivas para tareas de visión artificial, como en la clasificación y segmentación de imágenes, entre otras aplicaciones.

- *Mejorar el experimento:* en cuanto al número de participantes involucrados con el objetivo que, estadísticamente, los resultados fueran más concluyentes y en función de los escenarios simulados, pues si bien los participantes fueron conscientes de no estar ante una situación hostil real sino simulada al hacer el experimento, conlleva que se desvirtúe el principal escenario en el que el producto implementado debiera presentar su principal utilidad.
- *HMD militarizado:* el laboratorio de Visión Aplicada en conjunto con el departamento de I+D+i de Tecnologías Informáticas del CITEDEF han diseñado y desarrollado, mediante el uso de impresora 3D, cada uno de los componentes que forman parte del soporte de visión (gafa monocular Optinvent ORA-2¹⁶) y procesamiento (ODROID XU-4¹⁷ de escala más pequeña al ODROID XU-3) en un casco militar de kevlar. En su fase inicial de prueba se están ajustando cada pieza que conformaran el casco militar de RA (HMD). Se prevé próximamente contar con un HMD robusto y utilizable para realizar las primeras pruebas de concepto con los componentes tecnológicos correspondientes.



**Figura 10.1 – Prototipo de HMD diseñado y desarrollado en el CITEDEF
(fuente: Propia)**

¹⁶ Optinvent ORA-2: http://www.optinvent.com/our_products/ora-2/

¹⁷ ODROID-XU-4: https://www.hardkernel.com/main/products/prdt_info.php

Las funcionalidades detalladas anteriormente son las principales que se deberían llevar a cabo a fin de contar con un framework más robusto, más amplio, integrado a un HMD militarizado y que se adapte a diferentes condiciones ambientales.

Bibliografía

- [1] RDECOM, FUTURE SOLDIER 2030 Initiative, U. A. N. S. R. Center, Ed., USA: -, 2009.
- [2] D. W. Brown, «A Survey of Mobile Augmented Reality Technologies for Combat Identification Applications,» USA, 2012.
- [3] M. A. Livingston, L. J. Rosenblum, S. J. Julier, D. Brown, Y. Baillot, J. E. Swan II, J. L. Gabbard y D. Hix, An Augmented Reality System for Military Operations in Urban Terrain, -, Ed., December 2 -5, Orlando, Florida, USA: Proceedings of Interservice / Industry Training, Simulation & Education Conference (I/ITSEC), 2002.
- [4] Fundación Telefónica, Realidad Aumentada: una nueva lente para ver el mundo, Madrid, España: Editorial Ariel, S.A., 2011. Avda. Diagonal, 662-664. 08034 Barcelona (España), 2011.
- [5] J. Carmigniani y B. Furht, Augmented Reality: An Overview, Department of Computer and Electrical Engineering and Computer Sciences, Florida Atlantic University, Boca Raton, Florida, USA: B. Furht (ed.), Handbook of Augmented Reality, DOI 10.1007/978-1-4614-0064-6 1, © Springer Science+Business Media, LLC 2011, 2011.
- [6] M. A. Livingston, S. J. Julier y D. G. Brown, Situation Awareness for Teams of Dismounted Warfighters and Unmanned Vehicles, -, Ed., Naval Research Laboratory. Washington, DC, USA: SPIE Defense and Security Symposium. Enhanced and Synthetic Vision Conference., 2006.
- [7] A. R. Hevner, S. T. March, J. Park y S. Ram, «Design science in information systems research,» *Mis Quarterly*, 28(1), 75–105, 2004.
- [8] S. Tremblay, P. Jeuniaux, P. Romano, J. Lowe y R. Grenier, «A Multi-Perspective Approach to the Evaluation of a Portable Situation Awareness Support System in a Simulated Infantry Operation,» Miami Beach, FL, USA, 2011.
- [9] M. Gheisari y J. Irizarry, «Investigating Facility Managers' Decision Making Process through a Situation Awareness Approach,» *International Journal of Facility Management*. Vol 2, Nro. 1, 2011.

- [10] M. R. Endsley, C. A. Bolstad, D. G. Jones y J. M. Riley, «Situation Awareness Oriented Design: From User's Cognitive Requirements to Creating Effective Supporting Technologies,» Denver, Colorado, USA, 2003.
- [11] M. R. Endsley, «Design and evaluation for situation awareness enhancement,» Santa Mónica, CA, USA, 1988.
- [12] M. R. Endsley, «A taxonomy of situation awareness errors,» Aldershot, England, 1995.
- [13] J. G. Morrison, D. A. Kobus y C. M. Brown, «DARPA Improving Warfighter. Information Intake Under Stress. Augmented Cognition,» San Diego, CA, USA, 2006.
- [14] M. J. Zieniewicz, D. C. Johnson y J. D. Flatt, The Evolution of Army Wearable Computers, USA: PERVASIVE Computing, 2002.
- [15] J. Hicks, R. Flanagan, P. Dr. Petrov y A. Dr. Stoyen, «Eyekon: Distributed Augmented Reality for Soldier Teams,» 704 Edgewood Blvd, Papillion, NE, 68046, 2003.
- [16] T. Gjosaeter, «Interaction with mobile augmented reality,» Dissertation for the degree philosophiae doctor (PhD) at the University of Bergen, Oslo, 2015.
- [17] S. Sestito, S. Julier, M. Lanzagorta y L. Rosenblum, «Intelligent Filtering for Augmented Reality,» Sydney, Australia, 2000.
- [18] R. T. Azuma, «A Survey of Augmented Reality,» Teleoperators and Virtual Environments 6, 4, 355-385, -, 1997.
- [19] C. McDonald, Hand Interaction in Augmented Reality, Ottawa, Ontario, Canada: The Ottawa-Carleton Institute for Computer Science. School of Computer Science. Carleton University., 2003.
- [20] W. Le Roux, The use of Augmented Reality in Command and Control Situation Awareness, South Africa: Council for Scientific and Industrial Research, 2010.
- [21] F. M. Jiménez Montilla, Reconocimiento de objetos con Realidad Aumentada, Granada: Universidad de Granada, 2013.
- [22] M. Specht, S. Ternier y W. Greller, Dimensions of Mobile Augmented Reality for Learning: A First Inventory, Open University, the Netherlands: Journal of the Research Center for Educational Technology (RCET) 117. Vol. 7, No. 1, Spring 2011, 2011.

- [23] D. Wagner, «Handheld Augmented Reality,» PhD Thesis, Graz University of Technology Institute for Computer Graphics and Vision, Graz, Austria, 2007.
- [24] M. J. Abásolo Guerrero, C. Manresa Yee, R. Más Sansó y M. Vénere, Realidad virtual y realidad aumentada. Interfaces avanzadas., La Plata, Buenos Aires, Argentina: Editorial de la Universidad Nacional de La Plata (Edulp). 47 N° 380 / La Plata B1900AJP / Buenos Aires, Argentina, 2011.
- [25] A. MacWilliams, A Decentralized Adaptive Architecture for Ubiquitous Augmented Reality System, Institut für Informatik der Technischen Universität München, 2005.
- [26] A. MacWilliams, T. Reicher, G. Klinker y B. Bruegge, Design Patterns for Augmented Reality Systems, Institut für Informatik, Technische Universität München: In Proc. International Workshop Exploring the Design and Engineering of Mixed Reality Systems - MIXER 2004, 2004.
- [27] M. Kölsch, R. Bane y T. a. T. M. Höllerer, «Multimodal Interaction with a Wearable Augmented Reality,» IEEE Computer Graphics and Applications (Volume: 26, Issue: 3, May-June 2006), 2006.
- [28] D. Datcu y S. Lukosch, «Free-Hands Interaction in Augmented Reality,» New York, NY, USA , SUI '13 Proceedings of the 1st symposium on Spatial user interaction, 2013, pp. Pages 33-40.
- [29] P. Viola y M. Jones, «Rapid Object Detection using a Boosted Cascade of Simple Features,» de *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001*, Kauai, HI, USA, USA, 2001.
- [30] P. Tiefenbacher, S. Wichert, D. Merget y G. Rigoll, «Impact of Coordinate Systems on 3D Manipulations in Mobile Augmented Reality,» ICMI '14 Proceedings of the 16th International Conference on Multimodal Interactio, New York, NY, USA, 2014.
- [31] Y. Pang, M.L. Yuan, A.Y.C. Nee, S.K. On y Kamal Youcef-Toumi, A Markerless Registration Method for Augmented Reality based on Affine Properties, Australian Computer Society, Inc. Seventh Australasian User Interface Conference (AUIC2006), Hobart, Australia: Conferences in Research and Practice in Information Technology (CRPIT), Vol. 50. Wayne Piekarski, Ed. Reproduction for academic, 2006.
- [32] B. Furht, «Handbook of Augmented Reality,» Springer New York. Editor Affiliations: ID1. Dept. of Computer Science & Engineering, Florida Atlantic University, New York, USA, 2011.

- [33] S. You, U. Neumann y R. Azuma, «Orientation tracking for outdoor augmented reality registration,» IEEE Computer Graphics and Applications (Volume:19 , Issue: 6), Los Angeles, CA, USA, 2002.
- [34] B.-K. Seo , J. Park y J.-I. Park, «3-D visual tracking for mobile augmented reality applications,» Multimedia and Expo (ICME), 2011 IEEE International Conference on, Barcelona, Spain, 2011.
- [35] H. Kandil y A. Atwan, «A Comparative Study between SIFT- Particle and SURF-Particle Video Tracking Algorithms,» International Journal of Signal Processing, Image Processing and Pattern Recognition, Egypt, 2012.
- [36] A. Yilmaz, O. Javed y M. Shah, «Object Tracking: A Survey,» School of EECS, University of Central Florida, USA, 2006.
- [37] K. Pulli, A. Baksheev, K. Korniyakov y V. Eruhimov, «Real-time computer vision with OpenCV,» *Magazine Communications of the ACM*, vol. Volume 55 Issue 6, pp. Pages 61-69, 2012.
- [38] G. Bleser, «Towards Visual-Inertial SLAM for Mobile Augmented Reality,» (Thesis) Vom Fachbereich Informatik der Technischen Universität Kaiserslautern, 2009.
- [39] J. M. X. Natário Teixeira, «Analysis and evaluation of optimization techniques for tracking in augmented reality applications,» Recife, Brasil, 2013.
- [40] M. Guerrero, «A Comparative Study of Three Image Matcing Algorithms: Sift, Surf, and Fast,» Thesis of MASTER OF SCIENCE in Civil and Environmental Engineering. Utah State University., Logan, Utah, USA, 2011.
- [41] M. Maida, M. Preda y V. H. Le, «Markerless Tracking for Mobile Augmented Reality,» Signal and Image Processing Applications (ICSIPA), 2011 IEEE International Conference on, Kuala Lumpur, 2011.
- [42] D. van Krevelen y R. Poelman, «A Survey of Augmented Reality Technologies, Applications and Limitations,» The International Journal of Virtual Reality, 2010, 9(2):1-20, Delft University of Technology, The Netherlands, 2010.
- [43] R. Azuma, S. You y U. Neumann, «Hybrid Inertial and Vision Tracking for Augmented Reality Registration,» Proceeding, VR '99 Proceedings of the IEEE Virtual Reality, Washington, DC, USA, 1999.
- [44] K. Mikolajczyk y C. Schmid, «A Performance Evaluation of Local Descriptors,» IEEE Transactions on Pattern Analysis and Machine Intelligence (Volume:27 , Issue: 10), 2005.

- [45] T. Tuytelaars y K. Mikolajczyk, «Local Invariant Feature Detectors: A Survey,» Foundations and Trends in Computer Graphics and Vision Vol. 3, No. 3 (2007) 177–280, 2008.
- [46] M. Preda, Mixed and Augmented Reality Reference Model, C. P. M. P. Gerard J. Kim, Ed., Sapporo, JP: 10th AR Community Meeting, 2014.
- [47] T. Saarelainen y J. Jormakka, C4I2-Tools for the Future Battlefield Warriors., -: IEEE - Fifth International Conference on Digital Telecommunications, 2010.
- [48] M. A. Livingston, S. J. Julier, G. S. Schmidt y D. G. Brown, Mobile Augmented Reality: Applications and Human Factors Evaluations, A. I. T. C. 5. N. R. Laboratory, Ed., USA: -, 2006.
- [49] J. Timonen y J. Vankka, Enhancing Situational Awareness by Means of Information Integration of Sensor Networks and Visualization, S. 8756, Ed., Baltimore, Maryland, USA: SPIE 8756, 2013.
- [50] L. J. Rosenblum, S. K. Feiner, S. J. Julier, J. E. Swan II y M. A. Livingston, «The Development of Mobile Augmented Reality,» de *Expanding the Frontiers of Visual Analytics and Visualization*, -, Ed., Expanding the Frontiers of Visual Analytics and Visualization, pp 431-448, Springer London, 2012.
- [51] J. Juhnke, A. Kallish, D. Delaney, K. Dziedzic y R. Chou, «Tanagram Partners. Aiding Complex Decision Making through Augmented Reality: iARM, an Intelligent Augmented Reality Model,» USA, 2010.
- [52] Broad Agency Announcement, «Urban Leader Tactical Response, Awareness & Visualization (ULTRA-Vis),» Broad Agency Announcement (BAA), USA, 2008.
- [53] D. Roberts, A. Menozzi, J. Cook, T. S. S. Sherrilla, P. Russler, B. K. R. Clipp, E. Wenger, M. Bennett, J. Mauger, W. Church, H. Towles, S. MacCabe, J. Webb, J. Lupo, J.-M. Frahm, . E. Dunn, C. Leslie y G. Welch, Testing and evaluation of a wearable augmented reality system for natural outdoor environments, P. R. H. M. P. B. J. E. M. Peter L. Marasco, Ed., Head- and Helmet-Mounted Displays XVIII: Design and Applications: SPIE Vol. 8735, 2014.
- [54] R. E. Johnson, Components, Frameworks, Patterns, University of Illinois: Department of Computer Science, 1997.
- [55] D. C. Schmidt, Applying Patterns and Frameworks to Develop Object-Oriented Communication Software, D. o. C. Science, Ed., Washington University: Handbook of Programming Languages, 1997.

- [56] L. Fuentes, J. M. Troya y A. Vallecillo, *Desarrollo de Software Basado en Componentes*, E. I. C. Teatinos, Ed., Málaga, España: Dept. Lenguajes y Ciencias de la Computación, 2001.
- [57] S. Krakowiak, *Middleware Architecture with Patterns and Frameworks*, France: Project Sardes - INRIA Grenoble Rhône-Alpes , 2009.
- [58] D. Amin y S. Govilkar, «Comparative Study of Augmented Reality SDK's», *International Journal on Computational Sciences & Applications*, 5, 11-26., 2015.
- [59] F. Herpich, R. L. Martins Guarese y L. M. Rockenbach Tarouco, «A Comparative Analysis of Augmented Reality Frameworks Aimed at the Development of Educational Applications», *Creative Education*, Vol.08, Article ID:77994,19 pages , 2017.
- [60] V. Rautenbach y S. Coetzee, «Results of an Evaluation of Augmented Reality Mobile Development Frameworks for Addresses in Augmented Reality», *Conference: Conference: FOSS4G 2015, Seoul, Korea*, 2015.
- [61] C. Reynoso y N. Kicillof, «Estilos y Patrones en la Estrategia de Arquitectura de Microsoft», *Buenos Aires, Argentina*, 2004.
- [62] Duc Ngoc Tran, F. A. Hussin y M. Z. Yusoff, «Optimization of blood vessel detection in ret-ina images using multithreading and native code for portable devices», *Kuala Lumpur, Malaysia*, 2013.
- [63] C. Pujara, V. Veera, A. Kumar, N. Reddy y V. Tholath, «Real-time stereo video decoding and rendering on multi-core architecture», 2012.
- [64] R. Hofmann, H. Seichter y G. Reitmayr, «A GPGPU accelerated descriptor for mobile devices», 2012.
- [65] E. Rublee, V. Rabaud, K. Konolige y G. Bradski, «ORB: An efficient alternative to SIFT or SURF», 2011.
- [66] Yu-Hao Chang, Chi-Bang Kuan, Cheng-Yen Lin, Te-Feng Su, Chun-Ta Chen, Jyh-Shing Jang, Shang-Hong Lai y Jenq-Kuen Lee, «Support of Software Framework for Embedded Multi-core Systems with Android Environments», 2011.
- [67] Kwang-Ting Cheng, Xin Yang y Yi-Chu Wang, «Performance optimization of vision apps on mobile application processor», 2013.
- [68] D. A. Norman y Draper, S. W. , «User Centered Design», *illsdale, NJ: Lawrence Erlbaum*, 1986.

- [69] M. Maguire y N. Bevan, «User requirements analysis. A review of supporting methods,» Proceedings of IFIP 17th World Computer Congress, Montreal, Canada, 25-30. August 2002, p133-148. Kluwer Academic Publishers., Montreal, Canada, 2002.
- [70] C. Gutwin y S. Greenberg, «A Descriptive Framework of Workspace Awareness for Real-Time Groupware,» Computer Supported Cooperative Work, Kluwer Academic Press, 2001.
- [71] C. Ardito, M. T. Baldassarre, D. Caivano y R. Lanzilotti, «Integrating a SCRUM-based process with Human Centred Design: an Experience from an Action Research Study,» IEEE/ACM 5th International Workshop on Conducting Empirical Studies in Industry (CESI), 2017.
- [72] M. R. Endsley, «The Designer's Situation Awareness Toolkit: Support for User-Centered Design,» Article in Human Factors and Ergonomics Society Annual Meeting Proceedings, 2004.
- [73] M. R. Endsley y D. G. Jones, «Designing for Situation Awareness. An approach to User-Centered Design.,» CRC Press Taylor & Francis Group, 6000 Broken Sound Parkway NW, Suite 300. Boca Raton, FL 33487-2742, 2004.
- [74] Y.-G. J. A. G. H. C.-W. N. Jun Yang, «Evaluating bag-of-visual-words representations in scene classification,» MIR '07 Proceedings of the international workshop on Workshop on multimedia information retrieval, ACM New York, NY, USA, 2007, 2007.
- [75] T. M. K. Tuytelaars, «Local invariant feature detectors: a survey,» Journal Foundations and Trends® in Computer Graphics and Vision archive Volume 3 Issue 3, January 2008, Hanover, MA, USA, 2008, 2008.
- [76] D. Lowe, «Object recognition from local scale-invariant features,» The Proceedings of the Seventh IEEE International Conference on. vol. 2, pp. 1150-1157, Washington, DC, USA, 1999.
- [77] E. P. R. D. T. Rosten, «Faster and better: A machine learning approach to corner detection,» Intern Analysis and Machine Intelligence, IEEE Transactions on 32(1), 105-119, Washington, DC, USA, 2010.
- [78] R. Ortiz, «Freak: Fast retina keypoint,» Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). pp. 510-517. IEEE Computer Society, DC, USA, 2012.
- [79] I. Ojanen, «Introduction and integration of UCD and Scrum methodologies,» Aalto University. School of Science. Degree Programme in Computer Science and Engineering, Finlandia, 2016.

- [80] T. S. d. Silva, A. Martin, F. Maurer y M. Silveira, «User-Centered Design and Agile Methods: A Systematic Review,» Publisher: IEEE; Published in: 2011 Agile Conference, Salt Lake City, UT, USA, 2011.
- [81] M. R. Endsley, «Measurement of Situation Awareness in Dynamic Systems,» *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 37(1), pp. 65-84, 1995.
- [82] M. R. Endsley, «Situation Awareness Global Assessment Technique (SAGAT),» *Aerospace and Electronics Conference – Proceedings of the IEEE 1988 National*, 3, pp. 789-795, 1988.
- [83] R. S. Pierce, «Comparing Situation Awareness Measurement Techniques in a low fidelity air traffic control simulation,» In *Proceeding of the 26th International Congress of the Aeronautical Sciences (ICAS)*, pp. 1-8, 2008.
- [84] K. Finstad, «The Usability Metric for User Experience,» *Interacting with Computers*, pp. 323-327, 2010.
- [85] B. Bloom, «Taxonomy of educational objectives: the classification of educational goals,» Longmans, Green;, New York, 1956.
- [86] B. Kitchenham, *Procedures for Performing Systematic Reviews*, Keele University. Keele, Staffs ST5 5BG, UK: Technical Report TR/SE-0401. ISSN:1353-7776, 2004.
- [87] F. Muradas, «A Novel Framework for Requirements Elicitation in a Military Setting,» de *University of Oxford*, Oxford, UK, 2012.

III. Apéndices

Apéndice A. Cuadros de evaluación

Cuadro de evaluación de la CS de Endsley

Nivel	Pregunta	Puntaje por participante			
		P1	P2	P3	Total
Percepción	Q1				
	Q2				
	Q3				
	Q4				
Comprensión	Q5				
	Q6				
	Q7				
	Q8				
Proyección	Q9				
	Q10				
	Q11				

Cuadro de evaluación de la Experiencia del Usuario basado en el Cuestionario UMUX

Participante	1	2	3	4	5	6	TOTAL
Afirmación							
RAIOM cubre mis necesidades							
RAIOM es agradable en uso							
RAIOM es fácil de usar							
RAIOM no necesita ajustes							
TOTAL							

Apéndice B. Tareas realizadas por cada participante

Escenario: campo abierto, día soleado, visibilidad y luz óptimas. Se ubicó un objetivo a una distancia de 500 metros. Se cargó en el sistema posiciones (longitud y latitud) identificando el objetivo e información contextual hostil.

Lugar: CITEDEF (Instituto de Investigaciones Científicas y Técnicas para la Defensa). Villa Martelli, Provincia de Buenos Aires.

Grupo 1: grupo decisor / Jefe de comando (JPat).

- 1) Recorrer 150 metros hacia el objetivo
- 2) Determinar la ubicación geográfica.
- 3) Determinar la peligrosidad del entorno
- 4) Recorrer 100 metros. Realizar nuevamente las tareas 2 y 3
- 5) Recorrer 200 metros. Identificar distancia del objetivo. Realizar nuevamente las tareas 2 y 3
- 6) Recorrer 50 metros. Identificar vías de escape y ubicación de los elementos humanos y materiales

Grupo 2: grupo operativo / Comando (Pat).

- 1) Recorrer 150 metros hacia el objetivo
- 2) Determinar la ubicación geográfica.
- 3) Determinar la peligrosidad del entorno
- 4) Recorrer 100 metros. Realizar nuevamente las tareas 2 y 3
- 5) Recorrer 200 metros. Desplegar mapa virtual. Observar información provista.
- 6) Recorrer 50 metros. Reconocer objeto circundante. Informar al CII y al JPat

Apéndice C. Evaluaciones de performance de dispositivos móviles

Performance captura de vídeo y visualización

En este anexo se da cuenta de las actividades de medición de la performance relacionada con la adquisición de video tomando como base experimental las dos arquitecturas propuestas en la sección 8.3 - *Despliegues considerados*. En el experimento se evalúan los parámetros de performance correspondiente al tiempo de captura de video y visualización representado en cuadros por segundo *frame per second* (FPS) de las opciones analizadas. Para ello se ejecuta el método `onCameraFrame` que se implementó en el módulo Input para la adquisición de vídeo. El método `onCameraFrame` es un callback de la interface `CameraBridgeViewBase.CvCameraViewListener` que forma parte del conjunto de clases de OpenCV y su función es procesar los frames entrantes desde la cámara de video representando a la información como una matriz. La evaluación consistió en realizar cinco grupos de pruebas a las que llamaremos Test1, Test2, Test3, Test4 y Test5. Para cada grupo de prueba se ejecutó un total de 500 frames por cada dispositivo móvil. Luego se promedió el tiempo de adquisición de video y visualización de cada grupo por cada dispositivo móvil. Finalmente, se realizó un promedio general de los grupos de pruebas de cada dispositivo evaluado. Se debe considerar que la performance en gran medida se debe a la arquitectura de los dispositivos. En la Tabla 10.1 se detallan las arquitecturas de los dispositivos móviles evaluados:

Tabla 10.1 Arquitectura de los dispositivos móviles evaluados

Arquitectura	Dispositivo	CPU	Versión Android	Cámara
Opción 1: Gafa de RA	Epson Moverio BT 300	Quad Core Intel® Cherry Trail Atom™ x5 1.44GHz.	5.1	Built-in
Opción 2: Gafa emulada	Huawei RIO-L03 (G8) Android 5.1.1	Octa Core Qualcomm Snapdragon 616 MSM8939v2 4 x 1.5 GHz ARM Cortex-A53 4 x 1.2 GHz ARM Cortex-A53)	5.1.1	Built-in
	Samsung SM-G930F Android 6.0.1	Octa Core Exynos M1 Mongoose 4 x 2,6/2,3 GHz 4 x 1,5 GHz ARM Cortex-A53 (GTS)	6.0.1	Built-in
	Hardkernel ODROID-XU3 Android 5.1.1	Octa Core. Samsung Exynos 5422 4 x 2.0 GHz ARM Cortex-A15 4 x 1.4 GHz ARM Cortex-A7	5.1.1	USB (externa)

Con este resultado se puede calcular un valor aproximado para FPS con la siguiente fórmula:

$$fps = 1 / opfvalue$$

donde *opfvalue* es el valor en segundos del evento *onPreviewFrame*, antes calculado.

Se realizaron un total de 5 pruebas con cada opción de dispositivos. En la siguiente Tabla se muestran los resultados obtenidos:

Tabla 10.2 Comparación de performance correspondiente al tiempo de captura de video y visualización en los Dispositivos móviles utilizados en los Despliegues medido en milisegundos

Dispositivo	Delay en la captura de video (milisegundos) *	FPS
Epson Moverio BT 300	47,70	20,959
Hauwei G8	35,20	28,370
Samsung SM-G930F	55,00	18,172
ODROID XU3	50,40	19,829

* Promedio tomado de una muestra de 5 valores que arrojó la prueba

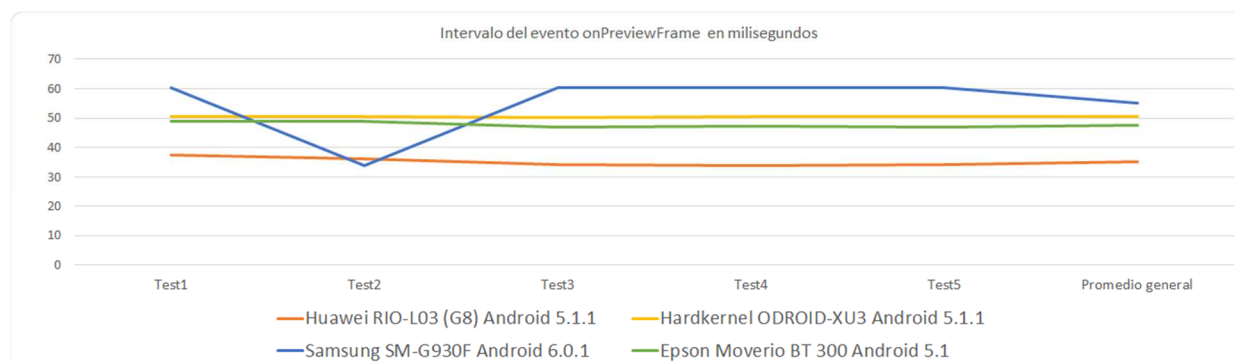


Figura 10.2 – Gráfico comparativo de performance entre los dispositivos medidos en segundos en el evento *onPreviewFrame*

A partir de las pruebas realizadas y teniendo en cuenta los promedios calculados, se puede concluir que el dispositivo que mejor se comportó en la obtención de frames es el Huawei G8. Las gafas Epson Moverio BT 300 tuvieron el segundo mejor comportamiento en la adquisición de vídeo, y se valora mejor su ergonomía.

Performance del proceso distribuido para el reconocimiento de objetos

Se tiene por objeto analizar la performance de todo el pipeline de reconocimiento de objetos en un ambiente distribuido. Se detalla que el proceso distribuido planteado está soportado por el diseño de una arquitectura Cliente / Servidor en donde del lado del Cliente se ejecutan los procesos menos intensivos; mientras que del lado del Servidor se implementó el módulo de *Vision* y algunas funcionalidades del módulo *Input* de entrada de datos ya que en este módulo se instauró el procesamiento de imágenes. La comunicación entre el *Cliente* (implementado en JAVA-Android) y el *Servidor* (implementado en JavaScript-Node.js) se realizó utilizando *Socket.io* como tecnología de comunicación entre estos. La evaluación consistió en capturar una imagen (frame) desde los dispositivos móviles analizados (*Cliente*) y enviar dicha imagen al mini board ODROID XU3 (*Servidor*) para su procesamiento en un ambiente distribuido. Se utilizó como enlace de comunicación inalámbrica un access point del tipo *Linksys Access Point Cisco* de doble banda. Seguidamente, el mini board procesa la imagen en el módulo *Vision* aplicando uno de los algoritmos propuestos para las pruebas. El resultado de aplicar dicho algoritmo a la imagen es devuelto al dispositivo móvil (*Cliente*) en forma de *keypoints* y la información del objeto reconocido. Finalmente, el módulo *Visualization* es el encargado de visualizar el resultado en la pantalla del dispositivo. Básicamente, la función de los algoritmos seleccionados del módulo *Vision* es detectar los puntos claves, armar los descriptores correspondientes, realizar los matchings (Brute Force), calcular las distancias de los pares de puntos claves, filtrar los puntos claves de baja calidad (RANSAC) y enviar el resultado al *Cliente* por medio de una conexión WiFi..

En la siguiente tabla se detallan las características principales de los componentes de hardware utilizados:

Tabla 10.3 Características principales de los dispositivos de hardware utilizados en la evaluación

Función	Tipo	Dispositivo	Software
Cliente	Opción 1: Gafa de RA	Epson Moverio BT 300	Android 5.1, OpenCV Manager, Socket.io y librerías (ver sección 7.3.2 - Capa: Librerías)
	Opción 2: Gafa emulada	Huawei RIO-L03 (G8)	Android 5.1.1, OpenCV Manager, Socket.io y librerías (ver sección 7.3.2 - Capa: Librerías).
Servidor	Mini board	ODROID XU3	Ubuntu Mate versión 16.04, Python versión 3.0, OpenCV 3.1 y Node.js.
Enlace de comunicación inalámbrico	Access point	Linksys Access Point Cisco Wap300n N300 Dual Band Wap 300	N/A

Tabla 10.4 Comparación de performance correspondiente al procesamiento distribuido medido en milisegundos

Servidor: ODROID XU3 (algoritmo)	Cliente Opción 1: Epson Moverio BT 300	Cliente Opción 2: Huawei RIO-L03 (G8)
ORB	17,60	44,70
FAST+FREAK	29,50	59,50
SURF	103,20	238,00

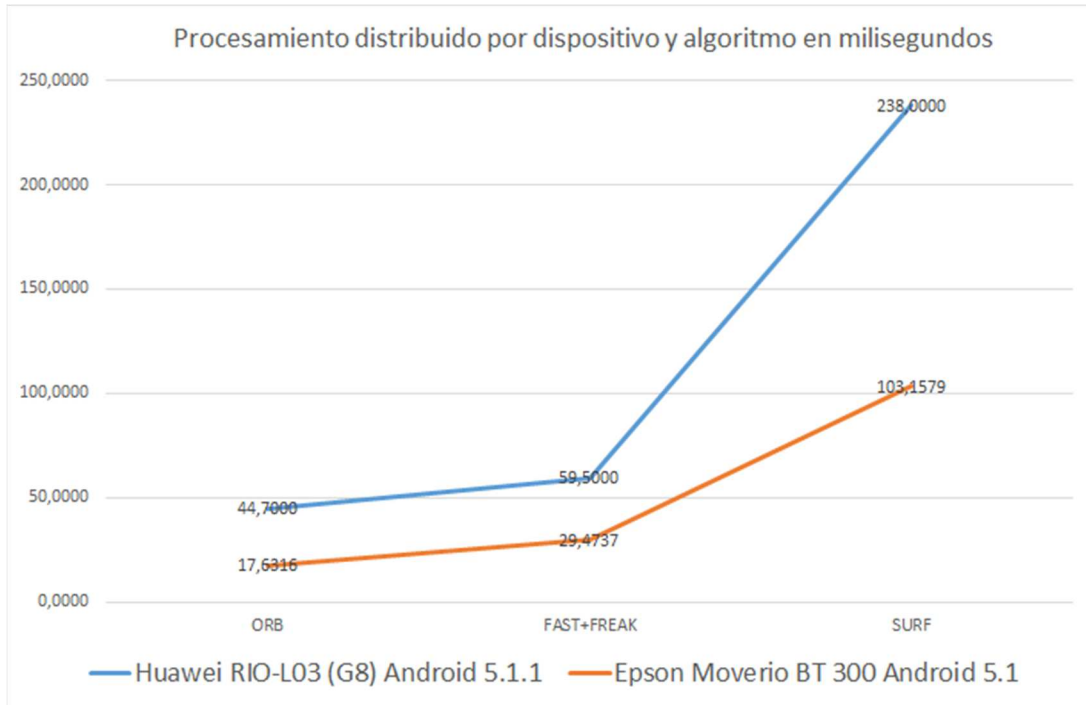


Figura 10.3 – Gráfico comparativo de performance del procesamiento distribuido evaluado en un dispositivo Huawei G8 y en una gafa Epson Moverio BT 300

Como conclusión se arriba a que las gafas Epson Moverio BT 300 tuvo mejor rendimiento que el dispositivo Hauwei G8 incluso teniendo éste mejores características de hardware (ver Tabla 10.1). Esto puede deberse a que el smartphone Huawei G8 posee aplicaciones en memoria ejecutándose simultáneamente lo que reduce la performance en comparación a las gafas Epson Moverio BT 300.

Comparación de performance en la ejecución del algoritmo ORB

Se realizó una prueba de performance entre Huawei RIO-L03 (G8) y mini board ODROID-XU3. La evaluación consistió en realizar veinte pruebas independientes de ejecución del algoritmo ORB en cada dispositivo reconociendo un objeto determinado. Luego se promedió el tiempo de ejecución del algoritmo ORB por cada dispositivo móvil. En la siguiente tabla se muestran los dispositivos evaluados y el tiempo promedio de procesamiento del algoritmo ORB de las veinte pruebas ejecutadas independientemente.

Tabla 10.5 Comparación de performance promedio correspondiente al procesamiento del algoritmo ORB en milisegundos

Dispositivo	Promedio de ejecución del algoritmo ORB en ms
Huawei RIO-L03 (G8)	516
ODROID-XU3	10

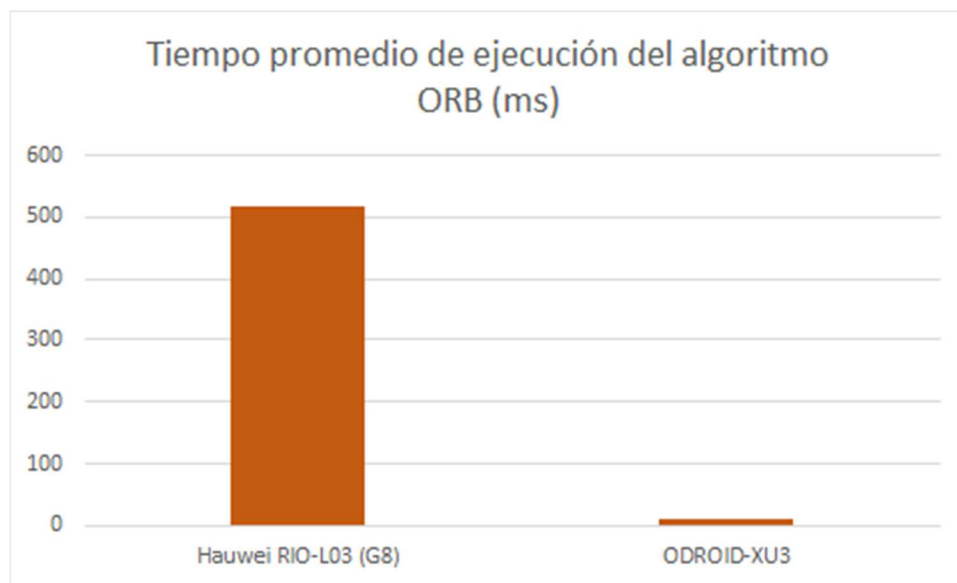


Figura 10.4 – Tiempo promedio de ejecución del algoritmo ORB en los dispositivos en milisegundos

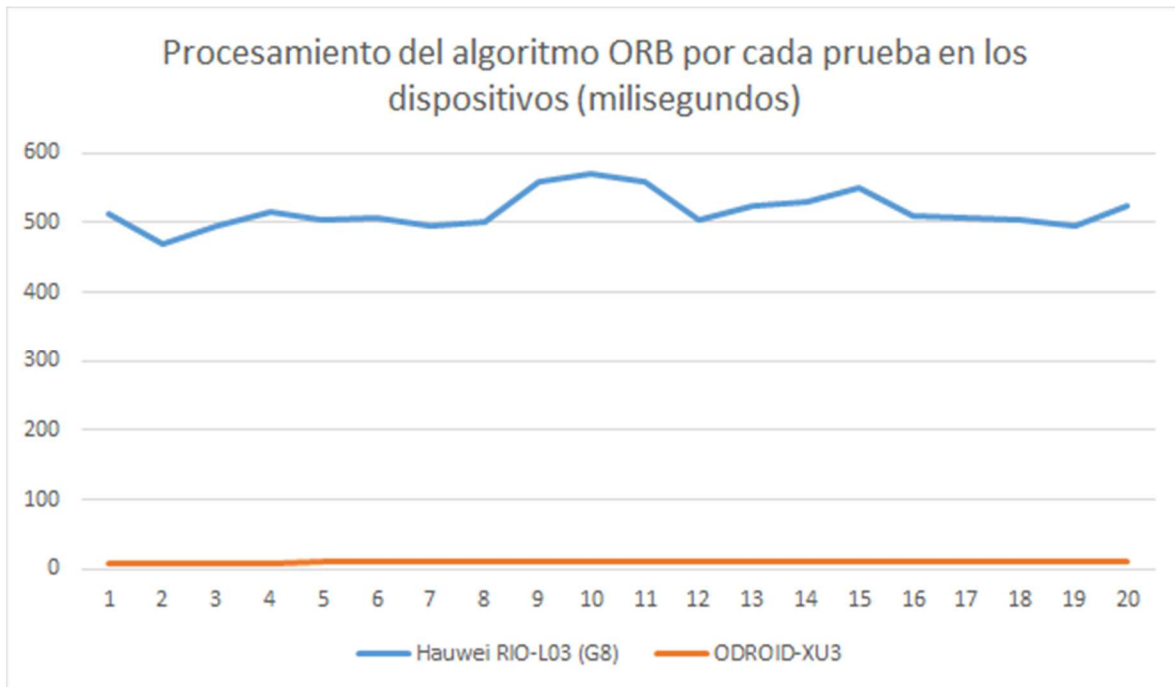


Figura 10.5 – Tiempo de ejecución del algoritmo ORB por cada una de las 20 pruebas realizadas en los dispositivos medido en milisegundos

Apéndice D. Diseño de prototipos utilizando historias de usuarios

En la propuesta de modelo tridimensional de CS (es decir, 3D Situational Awareness Model ó 3D-SAM) y tomando como base la Figura 6.5, las dos primeras dimensiones, fase y característica, son especialmente útiles para identificar los requisitos para los sistemas militares. La tercera dimensión, el ámbito (o alcance), cierra la brecha entre los requisitos y la solución de diseño basada en RA. En la Tabla 6.3 *Tabla de definición de CU por niveles del modelo de Endsley*, se identifican algunos ejemplos de los requerimientos principales en un dominio militar. Estos ejemplos pueden ser útiles en el desarrollo de sistemas militares, sin embargo, es difícil establecer una conexión entre los requisitos y el diseño para diseñadores nuevos. Por lo tanto, proponemos documentarlos como historias de usuarios. Las historias de usuarios son un concepto muy similar a los casos de uso tradicionales; son descripciones breves y simples de una característica contada desde la perspectiva de la persona que desea la nueva capacidad, generalmente un usuario o cliente del sistema. Para identificar una colección de historias de usuarios relevantes, se utilizó un marco para la obtención de requisitos en un entorno militar [87], donde el concepto de caso de uso fue cambiado por el término de historia del usuario. Al utilizar este marco, se identificaron, documentaron y crearon prototipos de un conjunto de historias de usuarios. Por ejemplo, un ejemplo de historia de usuario simplificada puede ser el siguiente texto: “Como soldado militar, quiero ver información sobre mis aliados y enemigos por medio de símbolos visuales; por ejemplo, se pueden mostrar puntos azules y rojos sobre el mapa del terreno”. Esta es una historia de usuario que se puede clasificar y documentar siguiendo el modelo tridimensional (fase = percepción, característica = presencia y alcance = visual) (ver el punto que se muestra como ejemplo en la Figura 6.5).

Con el modelo 3D-SAM, documentamos, identificamos y clasificamos los requisitos de SA y aportamos prototipos relacionados con las historias de usuarios del personal militar que contribuyen a su realización a nivel de análisis y de su posterior diseño.

A continuación se presentan tres ejemplos de historia de usuarios y el diseño de las soluciones siguiendo el modelo 3D-SAM.

Historia de Usuario 1: "Como comandante militar, quiero saber mi ubicación (dónde estoy) y mi dirección (dónde voy). En este sentido no solo sería útil conocer mi localización en base a coordenadas GPS sino que también resultaría adecuado que se identifiquen edificios, lugares o cualquier referencia que sea útil para conocer mi localización aproximada".

Esta es una historia de usuario que se puede clasificar y documentar siguiendo el modelo 3D-SAM:

(fase = *percepción*, característica = *dónde/localización* y modalidad = *visual*)

La Figura 10.4 muestra la representación gráfica del correspondiente punto del modelo 3D-SAM, el bosquejo del prototipo de la solución donde se muestra un radar de 360 grados, y una imagen del prototipo real implementado.

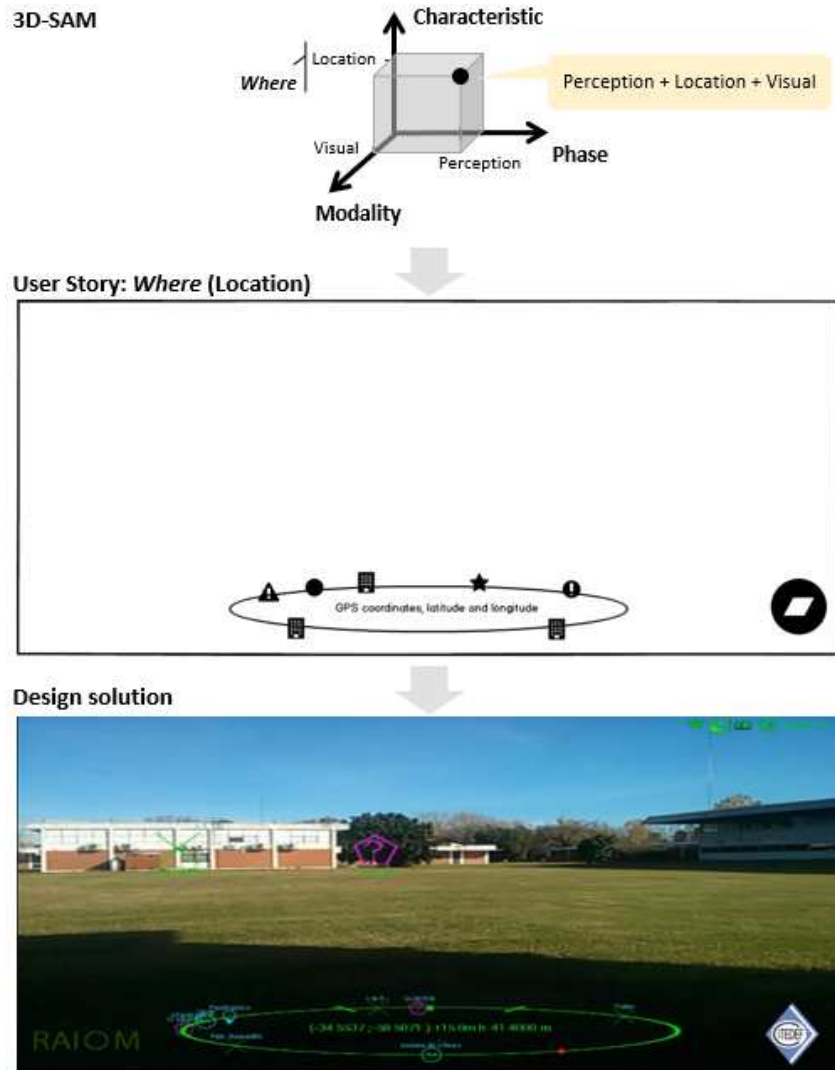


Figura 10.6 – Radar RAIOM de 360 grados. 3D-SAM, diseño e implementación de prototipos

Historia de Usuario 2: “Como comandante militar, quiero reconocer puntos de referencia para saber si este edificio es un lugar enemigo o aliado”.

Esta historia de usuario se puede clasificar y documentar siguiendo 3DSAM de la siguiente manera:

(fase = *percepción*, característica = *donde-ver*, y modalidad = *visual*)

La figura 10.5 muestra la representación gráfica del punto 3D-SAM correspondiente, el bosquejo del prototipo de la solución donde se pretende el reconocimiento de puntos de referencia y una imagen del prototipo real implementado.

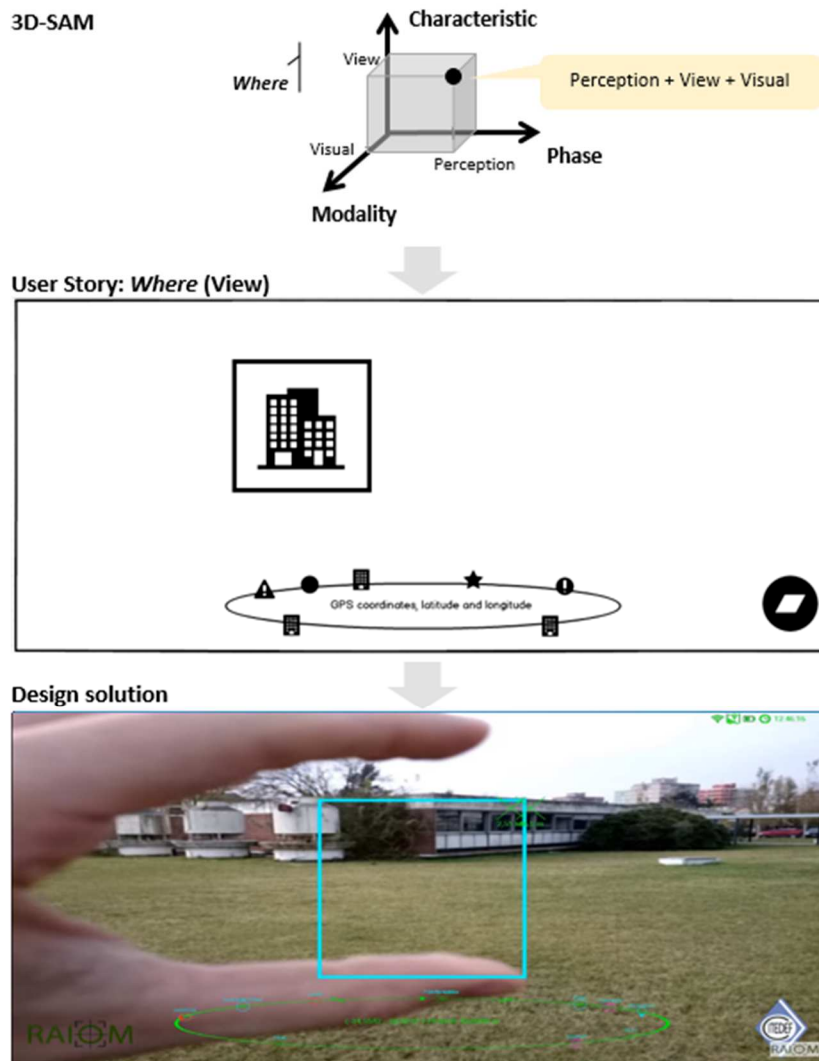


Figura 10.7 – Reconocimiento de puntos de referencia RAIOM. 3D-SAM, diseño e implementación de prototipos

Historia de Usuario 3: "Como comandante militar, quiero saber el lugar geográfico y quiero saber la ruta de escape en el mapa".

Esta historia de usuario puede clasificarse y documentarse siguiendo 3DSAM como sigue:

(fase = *percepción*, característica = *dónde-llegar* y modalidad = *visual*)

La Figura 10.6 muestra la representación gráfica del punto 3D-SAM correspondiente, el bosquejo del prototipo de la solución con un mapa interactivo y una imagen del prototipo real implementado.

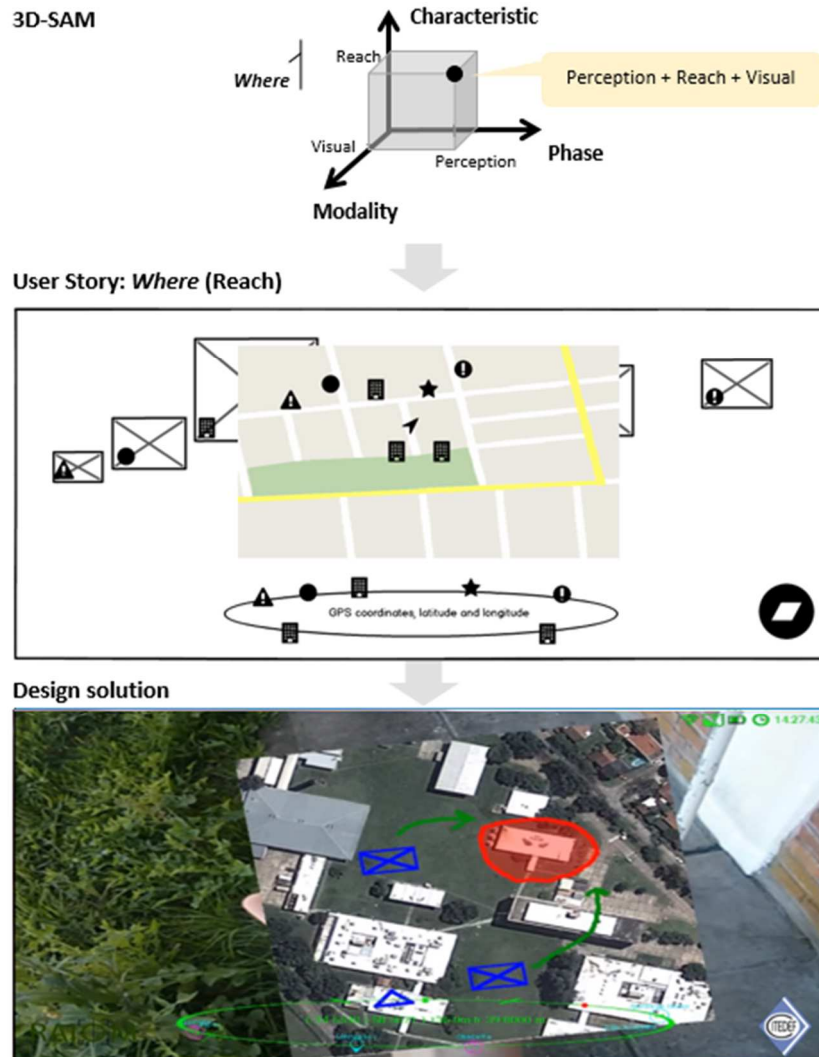


Figura 10.8 – Mapa interactivo de RAION. 3D-SAM, diseño e implementación de prototipos

Apéndice E. Videos demostrativos

1. **Prueba de Concepto #1:** Tracking de trayecto mediante arquitectura Cliente / Servidor

Link: <https://youtu.be/poSpWdzhLDU>

2. **Prueba de Concepto #2:** Reconocimiento gestual mediante la aplicación de técnicas como la envolvente convexa y tracking de puntos en la palma de la mano

Link: <https://youtu.be/gbZmGaML0tk>

3. **Prueba de Concepto #3:** Se integró en una aplicación desarrollada en Android el algoritmo de detección de características SIFT.

Link: <https://youtu.be/XHBjqv37s3k>

4. **Prueba de Concepto #4:** Integración del framework en una aplicación de visualización de alertas y amenazas mediante el uso de iconografía y simbología militar. Impresión en el campo de visión del operador de información contextual obtenida remotamente desde una unidad de comunicación.

Link: <https://youtu.be/OM-cIm6pHxI>

5. **Prueba de Concepto #5:** Manejo de menús a través de reconocimiento gestual.

Link: https://youtu.be/JtW_aMEquvs

6. **Prueba de Concepto #6:** Integración de varias funcionalidades en el framework.

Link: <https://youtu.be/hCSYqWnIs48>

7. **Pruebas con Usuarios**

Link:

https://drive.google.com/file/d/15ZRfN_XguB2OJS9t_ro805RSIEnrliA1/view?usp=sharing