

"SISTEMA AUTOMATIZADO PARA LA IDENTIFICACIÓN DE  
MACROINVERTEBRADOS ACUÁTICOS UTILIZANDO TÉCNICAS DE  
RECONOCIMIENTO DE IMÁGENES"

JULIO CESAR CASTILLO VERGEL

FUNDACION UNIVERSITARIA SAN MARTIN  
FACULTAD DE INGENIERIA  
BARRANQUILLA, COLOMBIA  
DICIEMBRE DE 2009

**“SISTEMA AUTOMATIZADO PARA LA IDENTIFICACIÓN DE  
MACROINVERTEBRADOS ACUÁTICOS UTILIZANDO TÉCNICAS DE  
RECONOCIMIENTO DE IMÁGENES”**

**JULIO CESAR CASTILLO VERGEL**

Proyecto de grado presentado como requisito para optar al título de “INGENIERO  
DE SISTEMAS”

Director: Ing. RENE RAMIREZ

**FUNDACION UNIVERSITARIA SAN MARTIN  
FACULTAD DE INGENIERIA  
BARRANQUILLA, COLOMBIA  
DICIEMBRE DE 2009**

Aprobado por el profesorado del Programa de Ingeniería de Sistemas en cumplimiento de los requisitos exigidos para optar al título de Ingeniero de Sistemas.



Ing. José Carmona  
Director Programa de Ingeniería de Sistemas

Ing. René Ramírez  
Director del Proyecto

Ruth Granados  
Asesora Metodológica

Bióloga Kelly Cuentas  
Asesora Científica

PÁGINA DE ACEPTACIÓN



---

DIRECTOR DE PROYECTO



---

JURADO CALIFICADOR



---

JURADO CALIFICADOR

Barranquilla, Diciembre de 2009

## CONTENIDO

	pág.
GUARDA .....	i
CUBIERTA .....	ii
PÁGINA DE ACEPTACIÓN.....	iv
DEDICATORIA .....	v
AGRADECIMIENTOS .....	vi
LISTA DE TABLAS .....	xiv
LISTA DE FIGURAS .....	xv
LISTA DE ANEXOS .....	xvii
GLOSARIO.....	xxiv
RESUMEN.....	xix
INTRODUCCIÓN .....	30
1. ESTADO DEL ARTE Y FUNDAMENTOS .....	34

1.1 PLANTEAMIENTO DEL PROBLEMA.....	34
1.2 FORMULACIÓN DEL PROBLEMA .....	36
1.3 OBJETIVOS.....	38
1.3.1 Objetivo general.....	38
1.3.2 Objetivos específicos: .....	38
1.4 JUSTIFICACIÓN .....	38
1.5 DELIMITACION.....	40
2. MARCO TEORICO.....	41
2.1 ANTECEDENTES Y REVISION CONOCIMIENTO DISPONIBLE .....	41
2.2 MARCO CONCEPTUAL O TEORICO.....	45
2.2.1 Conceptos procesamiento digital de imagenes .....	45
2.2.1.1 imagen digital .....	46
2.2.1.1.1 magenes vectoriales.....	46
2.2.1.1.2 Imágenes de mapa de bits .....	46
2.2.1.1.2.1 Imágenes de 1 bit por píxel .....	47

2.2.1.1.2.2 Imágenes de escala de grises (8 bits por píxel).....	48
2.2.1.1.2.3 Imágenes RGB (24 bits por píxel).....	48
2.2.1.1.2.4 Imágenes en color de 8 bits o menos .....	49
2.2.1.1.3. Formatos de imágenes digitales.....	49
2.2.1.1.3.1. Formato de imagen BMP.....	50
2.2.1.1.3.2. Formato de archivo GIF.....	50
2.2.1.1.3.3. Formato de archivo TIFF.....	50
2.2.2 Modelo general para el procesamiento de imágenes. ....	51
2.2.2.1 Adquisición de la imagen.....	52
2.2.2.2 Preprocesamiento. ....	52
2.2.2.3 Segmentación .....	52
2.2.2.4. Representación y descripción.....	53
2.2.2.5 Reconocimiento e interpretación .....	53
2.2.2.6. La base de conocimientos.....	54
2.2.3 Algebra de imágenes .....	55

2.3 DEFINICION DE CONCEPTOS Y VARIABLES .....	55
2.3.1 Variable dependiente .....	55
2.3.2 Variables independientes .....	56
2.3.3 Variables intervinientes. ....	56
2.4 HIPOTESIS.....	56
3. MARCO METODOLOGICO .....	57
3.1 TIPO DE INVESTIGACION O ESTUDIO.....	57
3.2 METODO .....	58
3.3 TECNICAS DE RECOLECCION DE LA INFORMACION.....	58
3.3.1 Metodos básicos. ....	59
3.3.1.1 La observación.....	59
3.3.1.2 La entrevista.....	60
3.3.2 Metodos hibridos.....	60
3.3.2.1 La investigación-acción .....	60
3.4 POBLACION OBJETIVO DE LA ENTREVISTA .....	61

3.5 JUSTIFICACION DE LA MUESTRA DE LA ENTREVISTA .....	61
3.6 ANALISIS DE LOS DATOS OBTENIDOS DE LAS ENTREVISTAS .....	61
3.6.1 Analisis e interpretacion de las entrevistas .....	61
3.6.1.1 Entrevista al decano de la facultad de ciencias básicas de la Universidad del Atlántico .....	61
3.6.1.2 Entrevista a catedrática del programa de biología de la facultad de ciencias básicas de la Universidad del Atlántico. ....	64
3.6.1.3 Entrevista a la especialista en identificación Taxónomica de macroinvertebrados acuáticos .....	66
3.7. PRESENTACION DE LOS RESULTADOS DE LAS ENTREVISTAS .....	68
3.8 CONCLUSION DE LAS ENTREVISTAS .....	68
3.9 UNIVERSO Y MUESTRA DE LOS MACROINVERTEBRADOS .....	69
4. PROPUESTA .....	71
4.1 SOLUCION .....	71
4.1.1 Adquisición de imagenes .....	72
4.1.1.1 Instrumento óptico .....	72
4.1.1.2 Camara digital .....	72

4.1.1.3 Conectividad de microscopio a computador.....	73
4.1.1.4 Visualización, captura y almacenamiento de imagenes.....	73
4.1.2 Segmentación.....	73
4.1.2.1 Extracción de características.....	75
4.1.2.1.1 Análisis de componenetes principales (PCA).....	76
4.1.2.1.2 Análisis discriminante lineal (LDA).....	76
4.1.3 Representación y descripción.....	77
4.1.3.1 Representación de imágenes por reducción de dimensiones.....	77
4.1.3.1.1 El eigespacio (PCA).....	79
4.1.4 Reconocimiento e interpretación.....	81
4.1.4.1 Aprendizaje supervisado.....	81
4.1.5 Base de conocimientos.....	82
4.1.6 Resultado del reconocimiento de imágenes.....	82
5. RESULTADOS.....	84
5.1 ENTRENAMIENTO DEL SISTEMA.....	84

5.2 RECONOCIMIENTO DEL MACROINVERTEBRADO .....	88
5.3 RENDIMIENTO Y EFICIENCIA DEL SISTEMA.....	89
5.3.1 Rendimiento.....	89
5.3.2 Eficiencia.....	91
CONCLUSIONES .....	92
RECOMENDACIONES .....	95
BIBLIOGRAFÍA.....	96
ANEXOS .....	101

## LISTA DE TABLAS

	pág.
TABLA 1. Listado de macroinvertebrados acuáticos identificados por el sistema automatizado.	69
TABLA 2. Resultados de la evaluación del clasificador	89
TABLA 3. Matriz de confusión de validación cruzada	90
FIGURA 4.	
FIGURA 5.	
FIGURA 6.	
FIGURA 7.	
FIGURA 8.	
FIGURA 9.	
FIGURA 10.	
FIGURA 11.	
FIGURA 12.	
FIGURA 13.	
FIGURA 14.	
FIGURA 15.	
FIGURA 16.	
FIGURA 17.	
FIGURA 18.	

## LISTA DE FIGURAS

	pág.
FIGURA 1. Imagen vector	47
FIGURA 2. Imágenes bitmap, imágenes raster, $I(r,c)$ .	47
FIGURA 3. Representación de una imagen mediante píxeles	48
FIGURA 4. Imagen monocromática, 1 bit por pixel	48
FIGURA 5. Imagen en escala de grises, 8 bits por pixel	49
FIGURA 6. Imagen RGB, 24 bits por pixel	49
FIGURA 7. Imagen en 256 colores	50
FIGURA 8. Pasos fundamentales del procesamiento de imágenes	52
FIGURA 9. Imagen digital, arreglo de $m \times n$ píxeles	54
FIGURA 10. Microscópio estereoscópico Leica EZ4 D	72
FIGURA 11. Conexión microscopio y computador	73
FIGURA 12. Combinación de iluminación del microscopio	74
FIGURA 13. Estructura sistemas de reconocimiento de patrones	75
FIGURA 14. Ejemplo de selección de características	78
FIGURA 15. Representación LDA	79
FIGURA 16. Imagen en formato JPG	84
FIGURA 17. Imagen en formato PGM	84
FIGURA 18. Representación tridimensional de las 30 imágenes de los macroinvertebrados	85

FIGURA 19. Representación bidimensional de las 30 imágenes de los macroinvertebrados	85
FIGURA 20. Proyección tridimensional de las imágenes generadoras en el subespacio PCA	86
FIGURA 21. Proyección bidimensional de las imágenes generadoras en el subespacio PCA	86
FIGURA 22. Proyección tridimensional de los individuos en el subespacio PCA	87
FIGURA 23. Proyección bidimensional de los individuos en el subespacio PCA	87

## LISTA DE ANEXOS

		pág.
ANEXO 1.	Proyeccion Macroinvertebrado 1 3D	101
ANEXO 2.	Proyección Macroinvertebrado 2 3D	101
ANEXO 3.	Proyeccion Macroinvertebrado 3 3D	102
ANEXO 4.	Proyección Macroinvertebrado 4 3D	102
ANEXO 5.	Proyeccion Macroinvertebrado 5 3D	103
ANEXO 6.	Proyección Macroinvertebrado 6 3D	103
ANEXO 7.	Proyeccion Macroinvertebrado 7 3D	104
ANEXO 8.	Proyección Macroinvertebrado 8 3D	104
ANEXO 9.	Proyeccion Macroinvertebrado 9 3D	105
ANEXO 10.	Proyección Macroinvertebrado 10 3D	105
ANEXO 11.	Proyeccion Macroinvertebrado 11 3D	106
ANEXO 12.	Proyección Macroinvertebrado 12 3D	106
ANEXO 13.	Proyeccion Macroinvertebrado 13 3D	107
ANEXO 14.	Proyección Macroinvertebrado 14 3D	107
ANEXO 15.	Proyeccion Macroinvertebrado 15 3D	108
ANEXO 16.	Proyección Macroinvertebrado 16 3D	108
ANEXO 17.	Proyeccion Macroinvertebrado 17 3D	109
ANEXO 18.	Proyección Macroinvertebrado 18 3D	109
ANEXO 19.	Proyeccion Macroinvertebrado 19 3D	110

ANEXO 20.	Proyección Macroinvertebrado 20 3D	110
ANEXO 21.	Proyeccion Macroinvertebrado 21 3D	111
ANEXO 22.	Proyección Macroinvertebrado 22 3D	111
ANEXO 23.	Proyeccion Macroinvertebrado 23 3D	112
ANEXO 24.	Proyección Macroinvertebrado 24 3D	112
ANEXO 25.	Proyeccion Macroinvertebrado 25 3D	113
ANEXO 26.	Proyección Macroinvertebrado 26 3D	113
ANEXO 27.	Proyeccion Macroinvertebrado 27 3D	114
ANEXO 28.	Proyección Macroinvertebrado 28 3D	114
ANEXO 29.	Proyeccion Macroinvertebrado 29 3D	115
ANEXO 30.	Proyección Macroinvertebrado 30 3D	115
ANEXO 31.	Proyeccion Macroinvertebrado 1 2D	116
ANEXO 32.	Proyección Macroinvertebrado 2 2D	116
ANEXO 33.	Proyeccion Macroinvertebrado 3 2D	117
ANEXO 34.	Proyección Macroinvertebrado 4 2D	117
ANEXO 35.	Proyeccion Macroinvertebrado 5 2D	118
ANEXO 36.	Proyección Macroinvertebrado 6 2D	118
ANEXO 37.	Proyeccion Macroinvertebrado 7 2D	119
ANEXO 38.	Proyección Macroinvertebrado 8 2D	119
ANEXO 39.	Proyeccion Macroinvertebrado 9 2D	120
ANEXO 40.	Proyección Macroinvertebrado 10 2D	120
ANEXO 41.	Proyeccion Macroinvertebrado 11 2D	121
ANEXO 42.	Proyección Macroinvertebrado 12 2D	121

ANEXO 43. Proyeccion Macroinvertebrado 13 2D	122
ANEXO 44. Proyección Macroinvertebrado 14 2D	122
ANEXO 45. Proyeccion Macroinvertebrado 15 2D	123
ANEXO 46. Proyección Macroinvertebrado 16 2D	123
ANEXO 47. Proyeccion Macroinvertebrado 17 2D	124
ANEXO 48. Proyección Macroinvertebrado 18 2D	124
ANEXO 49. Proyeccion Macroinvertebrado 19 2D	125
ANEXO 50. Proyección Macroinvertebrado 20 2D	125
ANEXO 51. Proyeccion Macroinvertebrado 21 2D	126
ANEXO 52. Proyección Macroinvertebrado 22 2D	126
ANEXO 53. Proyeccion Macroinvertebrado 23 2D	127
ANEXO 54. Proyección Macroinvertebrado 24 2D	127
ANEXO 55. Proyeccion Macroinvertebrado 25 2D	128
ANEXO 56. Proyección Macroinvertebrado 26 2D	128
ANEXO 57. Proyección Macroinvertebrado 27 2D	129
ANEXO 58. Proyección Macroinvertebrado 28 2D	129
ANEXO 59. Proyección Macroinvertebrado 29 2D	130
ANEXO 60. Proyección Macroinvertebrado 30 2D	130
ANEXO 61. Imagen en color de Ampullaridae, vista ventral y dorsal	131
ANEXO 62. Imagen en escala de grises de Ampullaridae, vista ventral y dorsal	131
ANEXO 63. Imagen en color de Ampullaridae Pomacea, vista ventral y dorsal	131

ANEXO 64. Imagen en escala de grises de Ampullaridae Pomacea vista ventral y dorsal	131
ANEXO 65. Imagen en color de Ceratopogonidae larva, vista ventral y dorsal	132
ANEXO 66. Imagen en escala de grises de Ceratopogonidae larva, vista ventral y dorsal	132
ANEXO 67. Imagen en color de Ceratopogonidae pupa, vista ventral y dorsal	132
ANEXO 68. Imagen en escala de grises de Ceratopogonidae pupa vista ventral y dorsal	132
ANEXO 69. Imagen en color de Chironomidae, vista ventral y dorsal	133
ANEXO 70. Imagen en escala de grises de Chironomidae, vista ventral y dorsal	133
ANEXO 71. Imagen en color de Conchostraca, vista ventral y dorsal	133
ANEXO 72. Imagen en escala de grises de Conchostraca, vista ventral y dorsal	133
ANEXO 73. Imagen en color de Hydrobidae, vista ventral y dorsal	134
ANEXO 74. Imagen en escala de grises de Hydrobidae, vista ventral y dorsal	134
ANEXO 75. Imagen en color de Hydrobidae littorina, vista ventral y dorsal	134
ANEXO 76. Imagen en escala de grises de Hydrobidae littorina vista ventral y dorsal	134
ANEXO 77. Imagen en color de Noteridae, vista ventral y dorsal	135
ANEXO 78. Imagen en escala de grises de Noteridae, vista ventral y dorsal	135
ANEXO 79. Imagen en color de Noteridae, vista ventral y dorsal	135
ANEXO 80. Imagen en escala de grises de Noteridae, vista ventral y dorsal	135

ANEXO 81. Imagen en color de Planorbidae, vista ventral y dorsal	136
ANEXO 82. Imagen en escala de grises de Planorbidae, vista ventral y dorsal	136
ANEXO 83. Imagen en color de Pleidae, vista ventral y dorsal	136
ANEXO 84. Imagen en escala de grises de Pleidae, vista ventral y dorsal	136
ANEXO 85. Imagen en color de Tubificidae, vista ventral y dorsal	137
ANEXO 86. Imagen en escala de grises de Tubificidae, vista ventral y dorsal	137
ANEXO 87. Imagen en color de Hydrophilidae, vista ventral y dorsal	137
ANEXO 88. Imagen en escala de grises de Hydrophilidae, vista ventral y dorsal	137
ANEXO 89. Imagen en color de Thiaridae, vista ventral y dorsal	138
ANEXO 90. Imagen en escala de grises de Thiaridae, vista ventral y dorsal	138
ANEXO 91. Imagen en color de Hydrachnidae, vista ventral y dorsal	138
ANEXO 92. Imagen en escala de grises de Hydrachnidae, vista ventral y dorsal	138
ANEXO 93. Imagen en color de Elmidae, vista ventral y dorsal	139
ANEXO 94. Imagen en escala de grises de Elmidae, vista ventral y dorsal	139
ANEXO 95. Imagen en color de Corixidae, vista ventral y dorsal	139
ANEXO 96. Imagen en escala de grises de Corixidae, vista ventral y dorsal	139
ANEXO 97. Imagen en color de Coenagrionidae, vista ventral y dorsal	140
ANEXO 98. Imagen en escala de grises de Coenagrionidae, vista ventral y dorsal	140

ANEXO 99. Imagen en color de Aeshnidae, vista ventral y dorsal	140
ANEXO 100. Imagen en escala de grises de Aeshnidae, vista ventral y dorsal	140
ANEXO 101. Imagen en color de Corixidae 2, vista ventral y dorsal	141
ANEXO 102. Imagen en escala de grises de Corixidae 2, vista ventral y dorsal	141
ANEXO 103. Imagen en color de Scirtidae, vista ventral y dorsal	141
ANEXO 104. Imagen en escala de grises de Scirtidae, vista ventral y dorsal	141
ANEXO 105. Imagen en color de Libellulidae, vista ventral y dorsal	142
ANEXO 106. Imagen en escala de grises de Libellulidae, vista ventral y dorsal	142
ANEXO 107. Imagen en color de Thiaridae 2, vista ventral y dorsal	142
ANEXO 108. Imagen en escala de grises de Thiaridae 2, vista ventral y dorsal	142
ANEXO 109. Imagen en color de Hydrachnidae 2, vista ventral y dorsal	143
ANEXO 110. Imagen en escala de grises de Hydrachnidae 2, vista ventral y dorsal	143
ANEXO 111. Imagen en color de Pyramideliidae, vista ventral y dorsal	143
ANEXO 112. Imagen en escala de grises de Pyramideliidae, vista ventral y dorsal	143
ANEXO 113. Imagen en color de Ancyliidae, vista ventral y dorsal	144
ANEXO 114. Imagen en escala de grises de Ancyliidae, vista ventral y dorsal	144
ANEXO 115. Imagen en color de Lampyridae, vista ventral y dorsal	144

ANEXO 116. Imagen en escala de grises de Lampyridae, vista ventral y dorsal	144
ANEXO 117. Imagen en color de Hydrophilidae larva, vista ventral y dorsal	145
ANEXO 118. Imagen en escala de grises de Hydrophilidae larva, vista ventral y dorsal	145
ANEXO 119. Imagen en color de Belostomatidae, vista ventral y dorsal	145
ANEXO 120. Imagen en escala de grises de Belostomatidae, vista ventral y dorsal	145

## GLOSARIO

**AGENTES RACIONALES:** un agente es un ente que actúa en un entorno, y responde a estímulos de éste. Este agente es racional cuando actúa de manera correcta según la información que posee. Para que un agente racional sea ideal debe realizar una acción que maximice la medida de éxito, a partir de la base de evidencias que obtiene dicha sucesión de percepciones, junto con el conocimiento que el agente posee.

**AGUAS EPICONTINENTALES:** se refiere a aquellas aguas que están situadas en el continente. Pueden ser dulces o salobres.

**ANÁLISIS DE COMPONENTES PRINCIPALES:** el análisis de componentes principales es una técnica matemática, utilizada para reducir la dimensionalidad de un conjunto de datos. Sirve para determinar el número de factores subyacentes explicativos tras un conjunto de datos que expliquen la variabilidad de dichos datos.

**ANÁLISIS DISCRIMINANTE LINEAL:** es un método estadístico para clasificar individuos u objetos en grupos exhaustivos y mutuamente excluyentes, previamente establecidos los grupos, llevando en cuenta los valores observados de un conjunto de variables independientes. Este análisis involucra la obtención de combinaciones lineales (o no lineales) de variables independientes que discriminarán entre grupos definidos a priori, de manera que los errores de mala clasificación deben ser mínimos.

**ANÁLISIS BACTERIOLÓGICO DEL AGUA:** es la fase final en la identificación de los agentes patógenos que puedan producir efectos fisiológicos adversos en cualquier organismo. La evaluación de la calidad bacteriológica de aguas proporciona seguridad para el consumo. Estos análisis son indicadores de la presencia de microorganismos que en un momento dado pueden afectar la salud pública y se brindan como una herramienta para evaluar los sistemas de tratamiento y la calidad de las aguas. Este análisis comprende: Número más Probable de Coliformes Totales, Número más Probable de Coliformes Fecales, Recuento Total en Placa de Microorganismos Heterótrofos, Evaluación de Remoción de Microorganismos etc.

**ANÁLISIS BIOLÓGICO DEL AGUA:** el estado biológico de un cuerpo de agua determina la calidad del mismo, es decir, establece el grado de contaminación, caracteriza la carga contaminante y la capacidad de autodepuración biológica. La caracterización biológica del agua comprende la identificación de organismos indicadores, frecuencia de aparición, índice de saprobiedad, Índice de Diversidad

de Especies (IDE) y Plancton: Fitoplancton, Determinación de biomasa y productividad, Zooplancton, Perifiton, Neuston, Macrofiton, macroinvertebrados, bentónico.

**ANÁLISIS FÍSICOQUÍMICOS DEL AGUA:** estos análisis permiten establecer si un agua es apta para el consumo, o si se requiere de algún tratamiento especial para potabilizarla. Es así como estos análisis representan el punto de partida para el diseño y ejecución de obras de Ingeniería, tendientes a optimizar la calidad del agua de abastecimiento. Comprenden los siguientes análisis: olor, color, transparencia, turbiedad, dureza, alcalinidad, conductividad, temperatura, fosfatos, nitratos, nitritos, CO<sub>2</sub>, oxígeno disuelto, pH, sólidos totales, DBO, DQO, entre otros.

**BIOINDICACIÓN:** es una técnica de evaluación ambiental que se ha consolidado como método para la detección y control de la toxicidad en un determinado ecosistema. Consiste en la utilización de organismos vivos, ya sean animales o vegetales, para medir y controlar la contaminación de un entorno determinado. Los indicadores biológicos reaccionan de una manera concreta ante ciertos agentes contaminantes, y se convierten en unos libros de instrucciones en los que los científicos leen el tipo de contaminante y el grado de toxicidad.

**CALIDAD DEL AGUA:** término referido a la composición del agua en la medida en que ésta es afectada por la concentración de sustancias producidas por procesos naturales y actividades humanas. Como tal, es un término neutral que no puede ser clasificado como bueno o malo sin hacer referencia al uso para el cual el agua es destinada.

**COLECCIÓN DE REFERENCIA BIOLÓGICA:** las colecciones son los depositarios de la biodiversidad, entendida como la riqueza, abundancia y variabilidad de todas las especies, comunidades y los procesos ecológicos y evolutivos que ocurren dentro de las mismas. Una colección permite que el estudio de todos estos aspectos sea más accesible y refutable. Por otra parte, las colecciones son archivos detallados de la vida en el planeta y por lo tanto sirven como modelos predictivos de lo que será la vida en el mismo.

**COMPUTACIÓN GRÁFICA:** es la rama de las ciencias de la computación que se encarga del estudio, diseño y trabajo del despliegue de imágenes en la pantalla de un computador a través de las herramientas proporcionadas por la física, la óptica, la térmica, la geometría, etc. La computación gráfica es un área de gran aplicabilidad en industrias como el diseño industrial, diseño gráfico, edición de video y entretenimiento.

**CLAVE DICOTÓMICA:** es una herramienta que permite identificar a los organismos. Una clave dicotómica se basa en definiciones de los caracteres morfológicos, macroscópicos o microscópicos; de ella parten dos soluciones

posibles, en función de si tienen o no tienen determinado carácter, repitiéndose el proceso de definiciones de características, hasta llegar al organismo en cuestión.

**DISEÑO DE SOFTWARE:** es definido como el proceso de definición de la arquitectura, componentes, interfaces y otras características de un sistema o componente que resulta de este proceso. El diseño de software es la tarea de representar en un lenguaje gráfico (usualmente UML) la solución de una problemática.

**ESPÉCIMEN:** la palabra espécimen significa muestra o ejemplar, sobre todo en la medida en que es representativa de una clase de objetos o entidades. En biología espécimen es aquel individuo o parte de un individuo que se toma como muestra, especialmente el que se considera representativo de los caracteres de la población a la que pertenece.

**ÍNDICE BIÓTICO:** estos índices se basan en el concepto de organismo indicador, suelen ser específicos para un tipo de contaminación y/o región geográfica. Permiten la valoración del estado ecológico de un ecosistema acuático afectado por un proceso de contaminación.

**IMAGEN:** en computación, representación visual de cosas en forma digital. Suele estar representadas por miles de píxeles (llamado raster) que, vistos en conjunto, forman una fotografía, un gráfico, etc. o pueden estar hechas por vectores.

**IMAGEN EN 3D:** una imagen en 3D, en definitiva, es una imagen en dos dimensiones que simula las tres dimensiones.

**IMAGEN RASTERIZADA:** es un fichero de datos que representa una matriz de píxeles (puntos de colores) denominada raster. En esta matriz el color de cada píxel es definido individualmente. Los gráficos rasterizados al ser ampliados comienzan a pixelizarse, o sea, se agrandan los elementos constituyentes del gráfico, y pierden calidad. Los gráficos rasterizados son útiles para imágenes fotográficas, las cuales no pueden ser representadas por vectores.

**INTERFAZ GRÁFICA:** hace referencia al conjunto de métodos para lograr interactividad entre un usuario y una computadora. Es el conjunto de formas y métodos que posibilitan la interacción de un sistema con los usuarios utilizando formas gráficas e imágenes. Con formas gráficas se refiere a botones, íconos, ventanas, fuentes, etc. los cuales representan funciones, acciones e información.

**LENGUAJE C++:** la definición "oficial" del lenguaje nos dice que C++ es un lenguaje de propósito general basado en el C, al que se han añadido nuevos tipos de datos, clases, plantillas, mecanismo de excepciones, sistema de espacios de nombres, funciones inline, sobrecarga de operadores, referencias, operadores para manejo de memoria persistente, y algunas utilidades adicionales de librería.

**C++** no es un lenguaje orientado a objetos puro (en el sentido en que puede serlo Java por ejemplo), además no nació como un ejercicio académico de diseño.

**LÉNTICO**: son cuerpos de agua cerrados que permanecen en un mismo lugar sin correr ni fluir, como los lagos, las lagunas, los esteros, o los pantanos.

**LÓTICO**: es un sistema de agua de corriente como los ríos, arroyos y manantiales.

**LIMNOLOGÍA**: es la rama de la ecología que estudia los ecosistemas acuáticos continentales (lagos, lagunas, ríos, charcas, marismas y estuarios), las interacciones entre los organismos acuáticos y su ambiente, que determinan su distribución y abundancia en dichos ecosistemas.

**MACROINVERTEBRADO**: organismos invertebrados con un tamaño superior a 0,5 mm, y que se hacen visibles al ojo humano. La gran mayoría (alrededor del 70 %) corresponden a grandes grupos de artrópodos (Crustacea, Insecta), donde formas larvianas de insectos son los más abundantes (e.g. Diptera, Coleoptera, Plecoptera, Ephemeroptera, Trichoptera). El resto lo componen pequeños moluscos, oligoquetos, sanguijuelas y planarias. Puedes encontrarlos bajo piedras, hojas, raíces y enterrados en la arena.

**MATLAB**: es un ambiente de computo, de alta ejecución numérica y de visualización. MATLAB integra el análisis numérico, cálculo de matrices, procesamiento de señales, y graficación, en un ambiente sencillo de utilizar, donde los problemas y sus soluciones son expresadas justamente como están escritas; a diferencia de la programación tradicional. El nombre de MATLAB surge de matrix laboratory, dado que el elemento básico de dato es una matriz, la cual no requiere de dimensionamiento, lo que permite resolver problemas en una fracción de tiempo.

**PATRON**: medida materializada, instrumento de medida, material de referencia o sistema de medida destinado a definir, realizar, conservar o reproducir una unidad o uno o varios valores de una magnitud para que sirvan de referencia.

**PROCESAMIENTO DE IMÁGENES**: es el conjunto de técnicas que permite modificar una imagen digital con el objetivo de mejorarla o extraer de ella información.

**SISTEMA AUTOMATIZADO**: es aquel que controla un ambiente recibiendo datos, procesándolos y devolviéndolos con la suficiente rapidez como para influir en dicho ambiente en ese momento.

**SISTEMA EXPERTO**: sistemas de computación basados en conocimientos cuyos componentes representan un enfoque cualitativo de la programación. Con los

sistemas expertos se busca una mejor calidad y rapidez en las respuestas dando así lugar a una mejora de la productividad del experto.

**SISTEMA HÍBRIDO:** sistema que emplea técnicas de distinta naturaleza.

**TAXONOMÍA:** ciencia de ordenar a los organismos en un sistema de clasificación compuesto por una jerarquía de taxones anidados. Ciencia de la clasificación de los seres vivos que tiene por objeto el establecimiento de grupos o categorías (taxones), en las que se reúnen organismos afines y semejantes. En rango decreciente, los taxones básicos son: Reino, Filo, Clase, Orden, Familia, Género y Especie.

**TAXÓN:** es un grupo de organismos emparentados, que en una clasificación dada han sido agrupados, asignándole al grupo un nombre en latín, una descripción, y un "tipo", de forma que el taxón de una especie es un espécimen o ejemplar concreto.

**TECNOLOGÍAS BIOINSPIRADAS:** nacen de la aplicación de conceptos de inspiración biológica al diseño de sistemas analíticos. El objetivo, en suma, es comprender e imitar la forma en que los sistemas biológicos aprenden y evolucionan. Para diseñar estos sistemas, además de utilizar la computación tradicional numérico-simbólica, se usan otras metodologías tales como las redes neuronales artificiales, la lógica difusa y la computación evolutiva. Por ello, este intento de emulación del funcionamiento de los seres vivos se debe apoyar en un entorno multidisciplinar que agrupa físicos, informáticos, electrónicos, microelectrónicos y áreas de la ingeniería, como la biomédica o la neuromórfica, y aspira a conseguir auténticos sistemas electrónicos dotados de sentidos artificiales que permitan facilitar un sinfín de tareas y resolver problemas hasta ahora no resueltos.

**VISIÓN POR COMPUTADORA:** estudia la estructura física tridimensional del mundo para el análisis automático de imágenes. Sin embargo, es necesaria la calidad en el uso de imágenes. Incluye muchas técnicas que son útiles para sí misma, ej., el procesamiento de imágenes (que se refiere) a la transformación, codificación y transmisión de las imágenes) y los patrones, de las cuales los patrones visuales son pero solo una instancia). sin embargo, la visión artificial incluye técnicas para la descripción útil de la forma y del volumen, para modelos geométricos, y para el llamado proceso cognoscitivo. Así, aunque la visión artificial se refiere ciertamente al procesamiento de imágenes, estas imágenes son solamente la materia prima de una ciencia mucho más amplia, la misma que se esfuerza en última instancia para emular las capacidades perceptivas del hombre y, quizás, para verter una luz sobre la manera por la cual él logra su interacción adaptativa y robusta con su ambiente.

## RESUMEN

Este trabajo de tesis muestra el diseño e implementación de un "SISTEMA AUTOMATIZADO PARA LA IDENTIFICACIÓN DE MACROINVERTEBRADOS ACUÁTICOS UTILIZANDO TÉCNICAS DE RECONOCIMIENTO DE IMÁGENES", el cual radica su importancia en el impacto que genera el uso de las herramientas informáticas en el diagnóstico de la calidad de los cuerpos de agua para determinar el impacto ambiental generado por las actividades humanas en su entorno.

El sistema basa su funcionamiento en la aplicación de los algoritmos de reconocimiento de patrones Análisis de componentes principales (PCA) y Clasificador de vecino más cercano (KNN); utilizados en otras áreas tales como identificación de rostros, reconocimiento de caracteres y otros, pero que no se habían utilizado en el reconocimiento de organismos biológicos utilizados en técnicas de biomonitoreo.

Este proyecto brindó una herramienta a los especialistas de análisis de calidad de agua, capaz de disminuir el tiempo de identificación de los macroinvertebrados y la clasificación de los especímenes recolectados para una mejor y más rápida recomendación a los organismos de control gubernamentales.

Palabras claves. PCA, KNN, Macroinvertebrados, Biomonitoreo, Reconocimiento, Patrones, Imágenes.

## INTRODUCCIÓN

Existe en el hombre un deseo profundo de poder reproducir la habilidad cognoscitiva por medios artificiales. La fascinación que la inteligencia como materia de estudio ha suscitado al género humano, puede verse reflejada en la aparición de una rama íntegra del estudio científico llamada "Inteligencia Artificial" o también estudio de la inteligencia.

Una de las habilidades más interesantes que se plantea al hablar de este tema es la habilidad del cerebro para poder reconocer patrones. Se denomina reconocimiento de patrones a la capacidad de poder interpretar una imagen compleja (una foto, lo que ve el ojo) y actuar en consecuencia. Las computadoras digitales fueron diseñadas a partir de una lógica binaria (de 2 valores 0 - 1 o Verdadero - Falso), lo cual, si bien facilitó su construcción, ha tenido como efecto una gran dificultad para procesar y reconocer imágenes, fotos, planos y dibujos.

Si bien la inteligencia artificial es un término relativamente nuevo, de acuerdo con John McCarthy (Conferencia de Dartmouth 1955), se aprecia con claridad que ramas de las ciencias exactas como la física, han buscado la manera de brindar interpretación a los acontecimientos que vemos a diario. Es así como la óptica se ha convertido en una herramienta fundamental en diversos campos de aplicación de las ciencias, dentro de las que podemos enumerar: las comunicaciones, sistemas de reconocimientos, seguimiento e identificación, sistemas de control de calidad, entre otros más.

El procesado de imágenes es una de las más importantes líneas de estudio en que se divide la óptica moderna. Se enmarca dentro del procesado de imágenes, cualquier manipulación de una imagen que permita modificarla de forma apropiada, o bien extraer determinada información de ella, bajo un tratamiento puramente digital, puramente óptico, o bien bajo sistemas híbridos opto-electrónicos.

Uno de los campos de aplicación del procesado de imágenes, es el campo de la biología, donde se ha utilizado el reconocimiento de organismos a través de instrumentos ópticos, tales como microscopios y estereoscopios, lo que ha permitido clasificar taxonómicamente los distintos especímenes observados. Sin embargo, esta labor se realiza usualmente de forma manual, donde el observador

se somete a largas y tediosas horas frente al lente del equipo óptico para cotejar lo visto con figuras o fotografías diseminadas en libros de referencia.

A fin de facilitar este proceso al observador, se presentó este trabajo de grado con el cual se pretendió tratar el procesado de imágenes en las ciencias biológicas, aplicado en el área de limnología, que se encarga de estudiar los ecosistemas acuáticos continentales (cuerpos de agua) y determinar la calidad de los mismos; a través de métodos biológicos de acuerdo con Ramón Margalef (Limnología 1983).

Uno de los métodos más utilizados mundialmente en opinión de (Alba -Tercedor & Sánchez - Ortega 1988, Johnson et al. 1992) es el de la bioindicación con macroinvertebrados acuáticos. El uso de estos organismos en la evaluación de la calidad de agua ha sido ampliamente utilizado (Cairns & Pratt, 1993; Hellawell, 1986), debido a que ofrecen numerosas ventajas como:

1. encontrarse en todos los ecosistemas acuáticos, por lo que favorecen los estudios comparativos,
2. Su naturaleza sedentaria, que permite un efectivo análisis espacial de los efectos de las perturbaciones (Slack et al., 1973; Hawkes, 1979; Penny, 1985; Hellawell, 1986; Abel 1989),
3. Presenta ventajas técnicas asociadas a los muestreos cuantitativos y análisis de las muestras, los que pueden ser realizados con equipos simples y baratos (Hawkes, 1979; Wiederholm, 1980; Suess, 1982; Penny, 1985; Hellawell, 1986);
4. La taxonomía de muchos grupos es bien estudiada (Hawkes, 1979; Suess, 1982, Hellawell, 1986, Abel, 1989) y
5. Existen numerosos métodos para el análisis de datos incluyendo índices bióticos y de diversidad, los cuales han sido ampliamente utilizados en biomonitoreos a nivel comunitario (Hellawell, 1986) y de respuestas individuales (Hawkes, 1979; Suess, 1982; Rosenberg et al., 1986).

Este método determina, que la presencia de una comunidad de estos organismos en un cuerpo de agua determinado, es un índice inequívoco de las condiciones que allí están prevaleciendo y que las fluctuaciones de contaminación que puedan presentarse, no son lo suficientemente fuertes como para provocar un cambio significativo en la misma.

Por ello con la colecta de los macroinvertebrados y su posterior identificación taxonómica se permite al personal interesado establecer el impacto que ha tenido dicho cuerpo acuático y así brindar un informe detallado, realizado en forma ágil y oportuna, por cuanto dicha labor se venía desarrollando de forma manual, surge la necesidad de contar con una herramienta capaz de agilizar la clasificación y

mantenimiento de la información, actualmente escasa y limitada, se hizo necesario, como innovación, diseñar y desarrollar este sistema.

El cual tuvo como objetivo fundamental, adquirir imágenes de los macroinvertebrados a través de un instrumento óptico, procesarla digitalmente e identificarla mediante una colección de referencia virtual que pueda ser usada para comparar las características morfológicas de los individuos en cualquier cantidad y combinación a fin de contribuir con los estudios limnológicos del país, resolver los problemas taxonómicos existentes o para mejorar nuestro conocimiento.

Para alcanzar el objetivo propuesto se trabajó bajo la metodología aplicada al diseño de software basado en la solución de problemas, mediante técnicas de ingeniería. Como resultado de la metodología empleada, se obtuvo la implementación de un sistema de manejo de información integral, fácil de usar, escalable, extensible, de costo accesible, compatible con estándares internacionales, eficiente, dirigido a una comunidad de usuarios de distinto nivel.

Por medio de este sistema se pudo contribuir, mediante la aplicación de operaciones automáticas, a reducir el tiempo de identificación, el desgaste del especialista y optimizar la entrega de resultados para la pronta toma de decisiones por parte de los entes reguladores del medio ambiente, así como la posibilidad de poder adicionar nuevos especímenes a la colección para ser utilizados en posteriores análisis.

Este proyecto se encuentra enmarcado en la línea de investigación de Inteligencia Artificial (área de investigación de las Ciencias de la Computación, cuyo objeto de estudio es la inteligencia, el término artificial se refiere al intento por crear mecanismos inteligentes en un medio no biológico.), y Computación Gráfica, se desarrolló en la ciudad de Barranquilla, en un periodo de tiempo comprendido entre Julio y Noviembre de 2009.

Para desarrollar este trabajo, se elaboraron 7 capítulos en los cuales se estructura la investigación realizada para alcanzar los objetivos propuestos.

En el primer capítulo, se esboza el planteamiento del problema con su respectiva formulación, se enmarcan los objetivos del proyecto así como la justificación y delimitación del mismo.

En el segundo capítulo, se tratan los antecedentes y las bases teóricas, se analizan los distintos proyectos e investigaciones en reconocimiento y tratamiento de imágenes y se esbozan los fundamentos teóricos de imágenes digitales y su tratamiento. Además se plantean las variables y la hipótesis planteada para esta investigación.

El tercer capítulo, hace relación al marco metodológico, en éste se trata el tipo de investigación, los métodos y técnicas de recolección de la información. Se presentan los análisis de las entrevistas realizadas así como la población objetivo de esta investigación con su justificación.

En el cuarto capítulo, se plantean las posibles soluciones y se detalla la opción escogida; además se describe la forma de la implementación enmarcada en los lineamientos del proyecto

El quinto capítulo detalla la presentación de los resultados obtenidos; describiendo cada uno dentro de las etapas del sistema tales como entrenamiento, reconocimiento, rendimiento y eficiencia del sistema. Se presentan redactados y explicados para facilitar el análisis e interpretación, los comentarios permiten argumentar las conclusiones y recomendaciones.

El sexto capítulo, destaca las conclusiones sobre los objetivos de la investigación, por medio de los cuales se evidencia la importancia de la investigación realizada, resaltando el alcance y novedad del proyecto augurando que ésta tendrá un impacto a mediano y largo plazo.

En el séptimo capítulo se destacan dos recomendaciones dadas por el autor para una investigación interdisciplinaria capaz de llevar la presente investigación a niveles internacionales.