
**Instituto Tecnológico de Costa Rica
Escuela de Matemática**

**Sistema automático de clasificación de abejas sin aguijón
(Apidae: Meliponini) basado en el contorno y
venación de sus alas.**

**Proyecto de investigación: # 5402-1701-0101
Informe Técnico**



Investigadores:

MSc. Geovanni Figueroa M. (Coordinador)

Lic. Pablo Prendas Rojas.

Abril – 2015

1 Contenido

1	Contenido	3
2	Índice de Figuras	5
3	Índice de Tablas	6
4	Datos generales	7
4.1	Nombre del proyecto	7
4.2	Código.....	7
4.3	Autores y direcciones.....	7
4.4	Colaboradores	7
4.5	Resumen.....	7
4.6	Palabras claves	8
5	Introducción	8
6	Marco Teórico	10
6.1.1	El sistema FABIS	11
6.1.2	El sistema ABIS	11
6.1.3	El Sistema MOBS.....	12
6.1.4	Sistemas que emplean información morfométrica	12
7	Metodología	13
7.1	Construcción de la colección de especímenes y la base de datos de fotográfica... 13	
7.1.1	Selección de la colección de abejas del CINAT.....	13
7.1.2	Colecta de especímenes	14
7.1.3	Protocolo para preservación de muestras	14
7.1.4	Ingreso en la colección	14
7.1.5	Etiquetado y Catalogación.....	15
7.1.6	Sistematización.....	15
7.1.7	Depósito y almacenamiento	15
7.1.8	Protocolo para la captura de fotografías	15
7.1.9	Fijación y montaje de las alas.....	16
7.1.10	Captura de Imágenes de las alas.....	16
7.2	Procesado de imágenes	19
7.2.1	Sustracción del fondo	20
7.2.2	Parametrización SIFT	21

7.2.3	Nuevo parámetro discriminador adicionado al método SIFT	25
7.3	Desarrollo del sistema de clasificación.....	27
8	Resultados	28
9	Discusión y conclusiones	30
10	Referencias	31

2 Índice de Figuras

Figura 1 Componentes del cuerpo de una abeja	9
Figura 2. Sitios de colecta de las abejas utilizadas en el proyecto	14
Figura 3 Imagen del ala anterior generada a partir del protocolo de captura de imágenes.	16
Figura 4 (a) muestra de fotografía original de la colección, (b) fotografía obtenida luego del filtrado por umbral, conversión a blanco y negro y operaciones morfológicas, (c) resultado del proceso de eliminación de fondo	20
Figura 5 Selección del ala de interés a partir de la cercanía del centro de la fotografía con el centro de masa del ala.....	21
Figura 6 Proceso de cálculo de diferencia de gaussianas en diferentes escalas	22
Figura 7 Ilustración del efecto del filtrado gaussiano y diferencia de gaussianas	23
Figura 8 Comparación de puntos con sus vecinos para discriminar valores mínimos y máximos.	23
Figura 9 Salida de la parametrización SITF para una foto con 4477 puntos claves.	24
Figura 10 Caso extremo de correspondencias equivocadas entre las especies	25
Figura 11 Efecto de la incorporación del nuevo discriminador en la correspondencias entre las especies <i>Cephalotrigona zexmeniae</i> y <i>Melipona beecheii</i>	26
Figura 12 Esquema del sistema de discriminación de especies de abejas.	27

3 Índice de Tablas

Tabla 1 Nombre científico y una fotografía de la colección usada en el proyecto.	17
Tabla 2 Totalidad de especies utilizadas en la etapa de entrenamiento.	28
Tabla 3 . Porción de la matriz cruzada discriminadora del algoritmo en etapa de entrenamiento.	29
Tabla 4 Porcentaje de éxito en la discriminación de abejas nativas costarricenses por género y por especie. .	30

4 Datos generales

4.1 Nombre del proyecto

Sistema automático de clasificación de abejas sin aguijón (Apidae: Meliponini) basado en el contorno y venación de sus alas.

4.2 Código

Proyecto de investigación # 5402-1701-0101

4.3 Autores y direcciones

MSc. Geovanni Figueroa M. (Coordinador), gfigueroa@itcr.ac.cr. Escuela de Matemática
Lic. Pablo Prendas Rojas. jpprendas@itcr.ac.cr. Escuela de Matemática.

4.4 Colaboradores

1. Melvin Ramírez Bogantes, Universidad Nacional de Costa Rica.
2. Carlos Travieso Gonzalez, Universidad de Las Palmas de Gran Canaria.
3. Ingrid Aguilar, CINAT, Universidad Nacional de Costa Rica.
4. Eduardo Herrera, CINAT, Universidad Nacional de Costa Rica.

4.5 Resumen

El objetivo del proyecto, desarrollar un sistema automático de clasificación taxonómica de abejas sin aguijón nativas de Costa Rica, a partir del procesado digital de imágenes de sus alas, fue satisfactoriamente alcanzado.

Métodos. La estrategia para la obtención y análisis de resultados fue trabajar con una muestra pequeña de imágenes en las que se calibraron los procesos de eliminación de fondo, parametrización SIFT y determinación de correspondencias; luego se entrenó el algoritmo con una muestra más grande y finalmente se validaron los cálculos en otra muestra de imágenes totalmente independiente a las utilizadas en las etapas anteriores.

Resultados. Se construyeron tres colecciones: una de especímenes de abejas montada en alfileres entomológicos, otra de alas derechas anterior y posterior montadas en un porta

objetos con un bálsamo de Canadá como medio de preservación inerte y una tercera colección de fotografías de estas placas. Se desarrolló un sistema de discriminación de especies de abejas sin aguijón costarricenses basado en la parametrización SIFT aplicada a fotografías de sus alas. Según la literatura sistemas de discriminación con alguna similitud al desarrollado, han alcanzado tasas de eficiencia por encima de 90%, pero han trabajado a nivel de género únicamente o bien con pocas especies. El sistema que se desarrolló tiene un porcentaje de éxito en la discriminación de especie de 86,5 mientras que la discriminación por género llegó al 97,52%, trabajando con 18 especies y 10 géneros.

Conclusiones. La investigación realizada nos permitió concluir que la información de las alas de las abejas es un buen discriminador en la identificación de la especie a la que pertenecen una abeja, sin embargo, un análisis detallado de las asignaciones de especies equivocadas nos permitió concluir que en algunas especies las alas son realmente muy parecidas, por lo cual, es necesario tomar en cuenta otros parámetros para su clasificación.

4.6 Palabras claves

Abejas sin aguijón, Clasificación de especies, SIFT, Procesamiento de imágenes

5 Introducción

Las abejas constituyen un grupo de insectos muy diverso e importante para la dinámica de bosques y agroecosistemas (Meléndez R., Meneses C., & Kevan G., 2013). Se conocen cerca de 20 000 especies, de las cuales la mayoría son de hábitos solitarios (Michener D., 2007). No obstante, la sobreexplotación de los bosques y recientemente el uso de pesticidas entre ellos Neonicotinoides, han provocado la disminución en el número de colonias de *Apis mellifera* (Apidae: Apini) y efectos adversos aún no cuantificados en otras especies, entre ellas las abejas nativas sin aguijón consideradas las abejas más importantes en la polinización de los bosques tropicales y subtropicales del mundo (Apidae: Meliponini) (Thompson, 2012), (Whitehorn, O'Connor, Walkers, & Goulson, 2012), (Blacquièrre, Smagghe, van Gestel, & Mommaerts, 2012). Éstos son algunos de los muchos efectos negativos de la degradación ambiental (Vaissière, Klein, Cane, & Steffan-Dewenter, 2006).

Además, se considera que uno de los problemas principales para la conservación de las abejas nativas en América Latina es la falta de información sobre su riqueza, diversidad, taxonomía, distribución, dinámica de sus poblaciones y el impacto de las actividades humanas en las especies de abejas (Freitas, y otros, 2009). Estos temas han sido poco estudiados en los trópicos americanos (Freitas, y otros, 2009). Considerando lo anterior se propone realizar actividades promotoras de la conservación de las abejas nativas como lo es la creación de colecciones y la identificación apropiada de las especies. Por ello, en Costa Rica y otros países de América Latina se atrae la atención del público a través de talleres, campañas de difusión sobre la importancia económica y ecológica de las abejas enfocándose en el grupo de “abejas nativas sin aguijón o meliponinos”.

Las abejas nativas sin aguijón, son un componente muy importante en los ecosistemas naturales de los trópicos y subtropicos del mundo ya que facilitan la polinización de la mayoría de las plantas cultivadas (Brown & Paxton, 2009). Son altamente sociales por lo que se les puede observar formando colonias permanentes. Estas abejas tienen un aguijón atrofiado el cual no pueden utilizar para la defensa. Se conocen más de 400 especies, de las cuales 300 se encuentran en los trópicos y subtropicos de América. En Costa Rica hay al menos 20 géneros y 59 especies de abejas sin aguijón (Camargo & Pedro, 2007), entre las que destaca una nueva especie descrita recientemente (*Nogueirapis costaricana*) (Ayala & Engel, 2014). La mayoría habitan por debajo de los 1500 msnm. Las especies más utilizadas en la práctica de la meliponicultura (crianza de abejas sin aguijón) en Costa Rica son *Tetragonisca angustula* y *Melipona beecheii* para un total de 16 especies domesticadas lo cual denota la importancia cultural y económica de estas abejas. Ellas han sido una parte fundamental de la cultura prehispánica y hoy en día están siendo utilizadas en una gran diversidad de proyectos enfocados en la economía familiar, conservación de los bosques y turismo comunitario. Siendo sus productos principales la miel, propóleos y polen. Estos productos son muy cotizados entre la población local por las propiedades antimicrobianas que las mismas poseen (Vaissière, Klein, Cane, & Steffan-Dewenter, 2006).

Las abejas son insectos con alas membranosas (Filo Artrópoda, Orden Hymenoptera). Su cuerpo se compone de tres partes: cabeza, tórax y abdomen (en las larvas, el abdomen y el tórax no siempre están diferenciados). El tórax está formado por tres partes, a saber: protórax, mesotórax y metatórax. Cada una de estas partes consta de un par de patas. En las abejas los dos pares de alas crecen en el tórax, dos en el mesotórax y otras dos en el metatórax. Las membranas superior e inferior de las alas cubren una red de tubos endurecidos, llamados nervios o venas, que les aporta rigidez a éstas (Figura 1). La disposición de las nervaduras o venación es característica de la mayoría de las especies de insectos y es muy utilizada por los entomólogos como base para su clasificación.

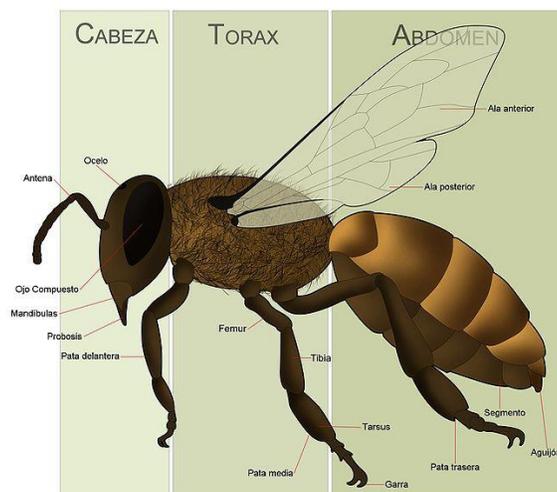


Figura 1 Componentes del cuerpo de una abeja

Algunas debilidades en cuanto a la taxonomía de las abejas en Costa Rica, en especial del grupo Meliponini, es la falta de colecciones entomológicas apropiadas que cubran inventarios de todo país, así como guías para su clasificación accesibles para el público. Se hace necesaria la búsqueda de sistemas fáciles y rápidos para evacuar dudas de identificación de especies, una necesidad cada día más evidente ya que los expertos por la lejanía y funciones propias de sus trabajos no pueden solventar (Arbuckle1, Schröder, Steinhage, & Wittmann, 2001). Un sistema digital de identificación de abejas de fácil uso y con resultados confiables sería una alternativa de solución al problema.

El conocimiento de las especies de abejas sin aguijón y su distribución en Costa Rica ayudará a desarrollar modelos de gestión y conservación de sus hábitats. En este sentido hay una necesidad latente de estudios sistemáticos que permitan una mejor comprensión de la diversidad de las abejas sin aguijón de Costa Rica.

Debido a esta necesidad y la importancia económica y social de estos insectos surgió el proyecto de investigación, con el objetivo de desarrollar un sistema automático de clasificación taxonómica de abejas sin aguijón nativas de Costa Rica, mediante el procesado digital de imágenes de sus alas.

6 Marco Teórico

La primera herramienta que utilizan los melitólogos para identificar una abeja son las guías de clasificación (llaves dicotómicas o claves dicotómicas). La identificación de especies y subespecies requiere enviar la muestra al melitólogo quién está entrenado para dicha identificación. Muchos de estos especialistas están retirados lo cual hará de esta faena más difícil en el corto y mediano plazo. Por ello, se hace necesario con ayuda de la tecnología buscar nuevas formas para simplificar este trabajo (Arbuckle1, Schröder, Steinhage, & Wittmann, 2001), (Hall, 2011).

Las guías para clasificación de abejas generalmente usan características fáciles de observar en el espécimen, la forma de las celdas en las alas, el largo de las alas tomando dos puntos de referencia en la misma y las proporciones entre estas distancias, por ejemplo. El número de celdas en las alas y su posición es similar entre algunos grupos, por ejemplo las abejas y avispas. Además, como en las abejas sin aguijón la venación es reducida, la clasificación hasta nivel de especie es complicada.

Tradicionalmente para la identificación de especies desconocidas se utiliza una clave dicotómica (clave taxonómica). La palabra “dicotómico” significa dividido en dos partes. Una clave dicotómica consiste en un número de pasos muy organizados en los que cada uno de los pasos a seguir tiene dos opciones mutuamente excluyentes, a veces conocidas como pareados y cada paso mejora la identificación. Para trabajar con una clave dicotómica se utilizan las distinciones más evidentes de la especie y con pares de afirmaciones contrapuestas por ejemplo "ojos labros" contra "ojos con pelos"). En <http://www.webbee.org.br/beetaxon/> se muestran ejemplos del uso de claves dicotómicas

para la clasificación de especies de abejas brasileñas.

Con el desarrollo de la tecnología y el crecimiento en la capacidad de procesamiento de las computadoras, se han propuesto sistemas de clasificación automáticos de abejas a partir de imágenes de sus alas muy eficientes. Existen estudios que demuestran que las alas contienen información importante para la identificación de insectos; ya que ellas varían mucho más entre especies que dentro de la misma especie e incluso las diferencias permiten discriminar entre subespecies (Tofilski, 2008), (Francoy, Silva, Nunes-Silva, Menezes, & Imperatriz-Fonseca, 2007). Las alas de los insectos son membranosas con venas visibles, lo cual permite la creación de un diagrama claro. En (Hall, 2011), (Breuker, Gibbs, Van Dongen, Merckx, & Van Dyck, 2010) se menciona que la clasificación de insectos basados en las características de la venación de las alas ha sido muy fructífera en los últimos años, pues el uso de indicadores morfométricos han cobrado relevancia en áreas biológicas.

6.1.1 El sistema FABIS

La primera propuesta de un sistema automático de clasificación de abejas se dio en la década de los 80's con la creación del Fast Africanized Bee Identification System (FABIS). No hay mucha información disponible sobre este sistema, sin embargo en (Francoy, y otros, 2008) se indica que el FABIS lograba clasificar de forma muy precisa abejas africanizadas. Para la discriminación de las abejas se necesitaba tomar 25 medidas morfométricas, esto requiere un proceso de entrenamiento y disponibilidad de varias horas para los cálculos. Esta situación provocó que el sistema se descontinuara, pero no así la iniciativa de automatizar la identificación de especies de abejas.

6.1.2 El sistema ABIS

Sin lugar a dudas, el aporte más reconocido al problema de automatizar la discriminación de abejas lo ha logrado el Automatic Bee Identification System (ABIS). ABIS fue propuesto en el año 1991 y tuvo seguimiento hasta 2008, lamentablemente después de este año no se le ha dado continuidad. La estrategia principal del ABIS consistió en utilizar la información del ala de forma similar al uso de la huella digital en humanos, para esto utilizaba una gran cantidad de información (300 características) para un posterior reconocimiento estadístico.

En particular tienen especial atención las celdas que se forman en el interior de las alas. Con la información de las líneas de venación e intersecciones de estas, el sistema genera hipótesis sobre la ubicación de las celdas. Las celdas creadas son comparadas con un árbol de hipótesis para ser clasificadas con Análisis de Discriminante Lineal, además la información de estas celdas se guarda para futuras guías. Los vectores de información obtenidos en el entrenamiento son complementados con las estimaciones del árbol de hipótesis y estos son clasificados mediante métodos que usan funciones kernel no lineales.

La información disponible sobre los resultados obtenidos con ABIS señalan un 99,8% de éxito en el reconocimiento a nivel de género. Sin embargo, es importante destacar que únicamente se trabajaron los géneros *Colletes*, *Andrea* y *Bombus*. A nivel de especie, en esos géneros, el éxito bajo a 95%.

Aunque la estrategia del sistema ABIS logró un alto porcentaje de acierto, se destacan entre sus limitaciones que utilizó muy pocas especies y géneros, demanda el entrenamiento del sistema con al menos 20 especímenes por clase (ya sea especie o género) y requiere que el usuario manualmente identifique algunas características.

Como consecuencia del éxito de este sistema han surgido, más recientemente, algunas propuestas que pretenden dilucidar las limitaciones reconocidas en ABIS. En 2001, se hace una publicación que destaca la precisión, portabilidad y herramientas utilizadas por ABIS pero se señalan como necesidades la creación de una aplicación de acceso público, monitoreo al sistema y la incorporación de información geográfica.

Un poco más reciente, en 2008, se plantea la posibilidad de combinar los resultados que genera el sistema ABIS con descriptores morfométricos presentes en las alas (Francoy, y otros, 2008). Esta propuesta impulsó la utilización de software complementario en los procesos de discriminación. En las fotografías, luego de ser procesadas con ABIS, se marcaban 19 intersecciones homólogas de venaciones con el uso del software tpsDig y estos puntos se alinearon y compararon con el uso de análisis de Procrustes. La combinación de ABIS con análisis morfométrico dio resultados con precisión de 98,05% en africanizadas y 80% con la inclusión de las otras 4 especies.

6.1.3 El Sistema MOBS

Por último, en 2011 se propone el MelittO Biotaxis System (MOBS) con la finalidad de identificar abejas vivas por las características anatómicas de sus alas (Hall C. , 2011). MOBS tuvo una eficiencia de 90% en la identificación de 16 especies agrupadas en 4 géneros. Lo realmente novedoso en MOBS es la reducción automática de las características utilizadas en la discriminación (menos de 20) en comparación con ABIS. En una primera etapa se extrae la venación de las alas (con el uso de umbrales y morfología matemática), se extraen y etiquetan las celdas (como en ABIS), extracción de cruces de venas y etiquetado de venas; en una etapa posterior se combinan técnicas (entre ellas PCA, SIFT y SURF) para seleccionar las características más importantes. En (Hall C. , 2011), se concluye que las cruces de venas y los centros de masas de las alas son excelentes puntos claves y advierte además que un pobre entrenamiento implicará problemas de clasificación.

6.1.4 Sistemas que emplean información morfométrica

Como complemento a los sistemas de clasificación anteriores y como una consecuencia natural del desarrollo tecnológico, en años recientes se ha intensificado la utilización de

información morfométrica en la identificación de especies y el uso de software específico que ayudan en esta tarea.

Se destaca la utilización de DrawWing, impulsado en 2007, para generar una máscara con el trazo de la venación y contorno del ala. En (Tofilski, 2008) se describe los requerimientos de este software, se promociona su uso como alternativa de identificación de especies y se deja en claro que es libre y modificable.

Otras iniciativas en la justificación del uso de software e información morfométrica se encuentran en (Hall, MacLeod, & Wardhana, 2014), (Yüksel & Tüzün, 2011). En (Yüksel & Tüzün, 2011) durante el 2011, se utilizaron los software TPSutil, TPSdig, CoordGen6f y CVAGen6o para apoyar el proceso desde la toma y procesado de la imagen, mediciones morfométricas hasta el análisis discriminatorio. Por otro lado, (Yüksel & Tüzün, 2011) en un estudio con alas de moscas utiliza el software tpsDig para la obtención de información morfométrica.

Hasta este punto se ha podido constatar que hay varios sistemas capaces de clasificar especies de abejas de forma eficiente, sin embargo también se pueden listar algunas limitaciones importantes. En primer lugar ninguno se ha planteado para la discriminación de abejas sin aguijón, se utilizaron pocas especies y géneros, los sistemas en su mayoría no son completamente automáticos (requieren intervención activa del usuario), algunos demandan muchas horas de entrenamiento y tal vez lo más relevante, están diseñados para clasificar un grupo de imágenes.

Por las razones expuestas, se ha considerado relevante la creación de un sistema de discriminación de abejas sin aguijón a partir de una fotografía del ala. Se propone un sistema de fácil utilización donde únicamente debe ingresarse una fotografía y este logre discriminar a qué especie pertenece esa ala.

7 Metodología

Para lograr los objetivos del proyecto este fue desarrollado por etapas.

7.1 Construcción de la colección de especímenes y la base de datos de fotográfica

Esta fue la primera etapa del proyecto, construir la colección de especímenes y la base de datos fotográfica que serían utilizadas a lo largo del proyecto de investigación. Para esto se llevaron a cabo las siguientes etapas.

7.1.1 Selección de la colección de abejas del CINAT

Parte de los especímenes utilizados fue tomada de la colección de abejas nativas sin aguijón del Centro de Investigaciones Apícolas Tropicales (CINAT). Se revisaron todos los especímenes para seleccionar aquellos que su ala derecha, anterior y posterior se encontraban

en perfectas condiciones. Las abejas seleccionadas para ser utilizadas en el proyecto fueron etiquetadas con un código especial y fueron organizadas en gavetas.

7.1.2 Colecta de especímenes

Para las especies que no se obtuvieron 15 individuos con sus alas en perfectas condiciones, fue necesario realizar visitas de campo para ubicar nidos, de donde se colectaron los especímenes necesarios para ser incluidos en la colección de abejas para el proyecto.

Las abejas se colectaron directamente de los nidos en diferentes zonas geográficas de Costa Rica (Figura 2. Sitios de colecta de las abejas utilizadas en el proyecto.), utilizando red entomológica y succionadores de plástico. Se utilizó una “cámara de muerte” con acetato de etilo para matar rápidamente y evitar daños en las alas. Las muestras se trasladaron al (CINAT) donde se realizó el proceso para la preservación de los ejemplares, el montaje y la identificación respectiva.

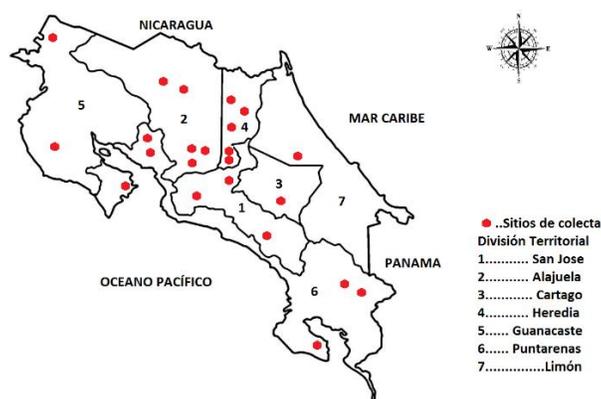


Figura 2. Sitios de colecta de las abejas utilizadas en el proyecto.

7.1.3 Protocolo para preservación de muestras

La creación de la base de datos requirió la preparación de tres colecciones; una de especímenes de abejas, otra de las alas derecha anterior y posterior montadas en un medio de preservación inerte y una tercera la colección de fotografías de las placas (portaobjetos con cubreobjetos y bálsamo de Canadá).

En esta etapa se debe asegurar la adecuada preservación de los ejemplares como lo es el ingreso de las muestras, cuarentena, etiquetado, catalogación, sistematización, identificación taxonómica y almacenamiento (Ramírez & Bernal, 2005).

7.1.4 Ingreso en la colección

Se hizo el montaje directo de las abejas con alfileres entomológicos (Insects pins BioQuip Black enamelled No. 1), se almacenaron para su secado, procesado, confección de etiquetas y posterior almacenamiento en las cajas entomológicas y gavetas destinadas para tal

colección. Los ejemplares presentan información que permite su identificación taxonómica (País, Provincia, Cantón, fecha, nombre del colector, muestra en nido o en flor). Con el cual se hace su registro en la colección. La identificación la llevó a cabo el MSc. E. Herrera del CINAT utilizando la clave de Ricardo Ayala para los géneros de ASA de Costa Rica (Ayala R., 2010 sin publicar).

7.1.5 Etiquetado y Catalogación

Se asignó un código (consecutivo) único para cada ejemplar. Las etiquetas son elaboradas en papel blanco tipo opalina, impresión láser, con una etiqueta de procedencia y una etiqueta con el código. La base de datos digital contiene los códigos y asociado a ellos toda la información de cada espécimen como lo es el nombre científico de la abeja. El mismo código se asigna a la muestra de ala y la fotografía de dicho ala. Las muestras pasan a una gaveta Cornell exclusiva para cuarentena. Se procede al montaje de las alas y abejas, y posteriormente a la captura de imágenes.

7.1.6 Sistematización

Toda la información del ejemplar se almacenó en una base de datos digital. Dicha información está al servicio de los investigadores asociados y del público con interés en estos registros.

7.1.7 Depósito y almacenamiento

Los ejemplares son colocados en cajas tipo “Cornell unit tray with polyethylene foam pinning bottom, 4-3/8 x 7-5/16””, de cartón las cuales se ingresan en las gavetas fabricadas de madera debidamente identificadas las cuales son depositadas en un cuarto de resguardo con las condiciones que debe tener una colección de insectos como lo es control de humedad, limpieza y temperatura para evitar su deterioro. Las muestras de alas en portaobjetos se resguardan en dicho cuarto en cajas de “Polystyrene Slide Box” de 100 Slides, con la rotulación respectiva. Las gavetas contienen pastilla de alcanfor. Y han sido desinfectadas previamente con alcohol al 75%.

7.1.8 Protocolo para la captura de fotografías

La fotografía debe cumplir con ciertos requerimientos de forma, color y texturas de forma que el ala se aprecie sin ruido y que la información que contenga sea únicamente del ala Figura 3 Imagen del ala anterior generada a partir del protocolo de captura de imágenes. En este proceso es importante que el ala pueda diferenciarse del fondo de la fotografía. Para lograr una buena calidad de las fotografías se estableció un protocolo de captura de fotografías (Herrera, y otros, 2013).

7.1.9 Fijación y montaje de las alas

Se extrajo el ala derecha anterior y posterior de las abejas montadas en alfileres entomológicos, se cortó en la base del mesotórax con navajilla y tijeras de disección de punta fina para facilitar el desprendimiento y no dañar las alas. Se quitó el residuo de la base del ala para su montaje. En un portaobjetos se colocó una gota de bálsamo de Canadá. Para suavizar el bálsamo se utilizó una lámpara de 50 w con el calor la muestra queda inmersa en el sustrato sin ser dañada. Se colocó el cubre objetos con presión moderada para extraer las burbujas y evitar el desplazamiento del ala. Se dejaron reposar sobre papel toalla por 24 horas antes de ser fotografiadas y almacenadas. La colección de alas comprende un total de 879 placas.

7.1.10 Captura de Imágenes de las alas

Se utilizó para la toma de imágenes un estereoscopio Olympus modelo SZX16 con una cámara integrada marca Olympus modelo DP72, la fuente de luz circular intralux 4000-1 con una intensidad entre 1400 y 1600fc, dependiendo de la muestra. La resolución en el milimétrico fue de 2x con una apertura focal de 1/4. Para la captura de imágenes se utilizó un sistema TWAIN con exposición manual, tiempo de exposición 60%, sensibilidad ISO 200 y una resolución Live 1360 x 1024 y un snap 4140 x 3096.

Las fotografías generadas fueron almacenadas en una base de datos, agrupadas por especie. La colección de fotografías comprende un total de 1758 fotografías.



Figura 3 Imagen del ala anterior generada a partir del protocolo de captura de imágenes.

En la figura 3 se muestra el nombre científico y una fotografía de las especies de la colección que fue utilizada en el proyecto.

Tabla 1 Nombre científico y una fotografía de la colección usada en el proyecto.

No	Especie	Fotografía base de datos y colección
1	<i>Cephalotrigona zexmeniae</i>	
2	<i>Dolichotrigona schulthessi</i>	
3	<i>Geotrigona lutzi</i>	
4	<i>Lestremielitta danuncia</i>	
5	<i>Melipona beecheii</i>	
6	<i>Scaptotrigona pectoralis</i>	
7	<i>Melipona costaricensis</i>	
8	<i>Nannotrigona mellaria</i>	
9	<i>Nannotrigona perilampoides</i>	

		
10	<i>Oxytrigona mellicolor</i>	
11	<i>Partamona orizabaensis</i>	
12	<i>Plebeia frontalis</i>	
13	<i>Plebeia jatiformis</i>	
14	<i>Paratrigona ornaticeps</i>	
15	<i>Scaptotrigona luteipennis</i>	
16	<i>Trigona nigerrima</i>	
17	<i>Tetragona zieglerei</i>	
18	<i>Scaptotrigona subobscuripennis</i>	

		
19	<i>Trigona fulviventris</i>	
20	<i>Trigona fuscipennis</i>	
21	<i>Trigona silvestriana</i>	
22	<i>Trigona muzoensis</i>	

En total se procesó material correspondiente a un 90% de los géneros de abejas nativas sin aguijón presentes en Costa Rica (20 géneros en Costa Rica) y un 73% de las especies reportadas (59 especies reportadas en Costa Rica) (Camargo & Pedro, 2007), (Gonzalez, 2000) .

7.2 Procesado de imágenes

Una vez que se construyó la base de datos de imágenes de las alas inicia el preprocesado de las mismas.

7.2.1 Sustracción del fondo

Una de las tareas más comunes en las investigaciones relacionadas con procesado de imágenes es la sustracción del fondo. Con mucha frecuencia en las fotografías aparecen elementos que no son importantes para los fines que se tomaron, por ejemplo, para los objetivos del proyecto solo interesaban los pixeles que corresponden al ala de la abeja. Dado el tamaño de las fotografías y el hecho que las alas corresponden a menos de la mitad de los pixeles totales, la discriminación de pixeles de fondo se traduce en un sustancial ahorro de procesado de información.

Existen muy diversas técnicas para distinguir cuando un pixel pertenece al fondo de la imagen y cuando no (Prendas & Ramírez, 2011), (González & Woods, 2012), (Nixon, 2012) por las características de las fotografías de las abejas se utilizaron histogramas, discriminación por umbrales y operadores morfológicos básicamente.

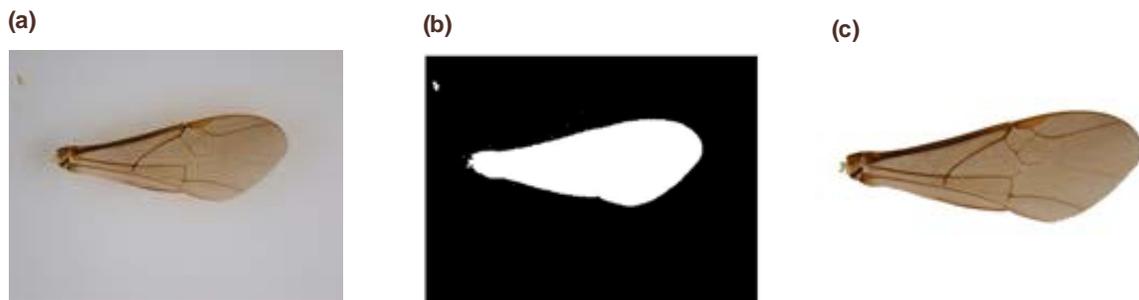


Figura 4 (a) muestra de fotografía original de la colección, (b) fotografía obtenida luego del filtrado por umbral, conversión a blanco y negro y operaciones morfológicas, (c) resultado del proceso de eliminación de fondo

En la colección de fotografías de abejas se logró distinguir esencialmente dos fondos; uno con un tono gris y otro con presencia de azul. Esto exigió que el algoritmo, como primera tarea, determinara cual tipo de fondo presentaba la foto, pues de esta primera decisión dependería la aplicación del método Otsu y los ajustes de umbrales. Para este fin se utilizó el espacio de color HSV y una cota, calculada empíricamente, para discriminar cuál de los fondos presentaba la foto. Una vez identificado el tipo de fondo, se implementó un filtrado mediante umbralización (Otsu) ajustada experimentalmente, para determinar qué pixeles se convertirían en blancos y cuáles en negros. En la Figura 4 (a) muestra de fotografía original de la colección, (b) fotografía obtenida luego del filtrado por umbral, conversión a blanco y negro y operaciones morfológicas, (c) resultado del proceso de eliminación de fondo, la región en blanco de la segunda fotografía corresponde a lo que se identificó como ala y la última imagen corresponde a la recuperación del color original en los puntos con asignación 'blanco'.

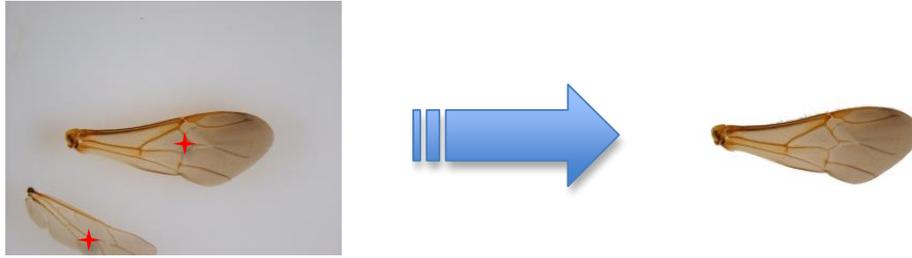


Figura 5 Selección del ala de interés a partir de la cercanía del centro de la fotografía con el centro de masa del ala

Una problemática adicional que se presentó en el proceso de eliminación de fondo fueron las fotografías con presencia de dos alas. Durante el protocolo de toma de fotografías, se decidió centrar el ala de interés para que este fuese el elemento discriminador en las que hubiese dos o más alas.

Para la implementación del algoritmo de discriminación de especies de abejas, las fotos solo deben contener un ala, pues si una foto tuviese dos alas podrían darse correspondencias incorrectas y esto generar asignaciones de especies equivocadas. Para solventar esta problemática, se incluyó en el algoritmo de eliminación de fondo el cálculo de los centros de masas de todas las regiones blancas que resultasen (si hay más de un ala habrá más de una región blanca) y medir las distancias de estos puntos hasta el centro de la imagen, la distancia más corta discriminará la región blanca que corresponde con el ala de interés. Una ilustración de esta estrategia se muestra en la Figura 5 Selección del ala de interés a partir de la cercanía del centro de la fotografía con el centro de masa del ala.

7.2.2 Parametrización SIFT

El Scale Invariant Feature Transform (SIFT) es un método propuesto por David Lowe en 1999, cuya idea básica es la transformación de una imagen a una representación compuesta de “puntos de interés”. Esta parametrización ha sido utilizada en muy diversas áreas como seguridad, reconocimiento facial, identificación de objetos, entre otras (Tofilski, 2008), (Francoy, Silva, Nunes-Silva, Menezes, & Imperatriz-Fonseca, 2007). Sin embargo, y esto debe remarcar, su uso se ha difundido mayormente a reconocimiento del mismo individuo o bien de objetos idénticos (Francoy, Silva, Nunes-Silva, Menezes, & Imperatriz-Fonseca, 2007), (Breuker, Gibbs, Van Dongen, Merckx, & Van Dyck, 2010), por lo que la implementación del SIFT para discriminación de especies de insectos es novedoso y debe hacerse con especial cuidado por las variantes naturales de los individuos aunque sean de la misma especie.

En ese sentido, la escogencia de la parametrización SIFT se fundamenta en la mayor cantidad de puntos claves que genera y el menor coste de procesado, en comparación con SURF y ASIFT por ejemplo (Breuker, Gibbs, Van Dongen, Merckx, & Van Dyck, 2010), (Hall, MacLeod, & Wardhana, 2014). Dado que las fotografías no corresponden al mismo individuo

se consideró importante generar una cantidad grande de puntos claves que describiesen cada imagen y luego ser exhaustivos en el pareo de estos puntos.

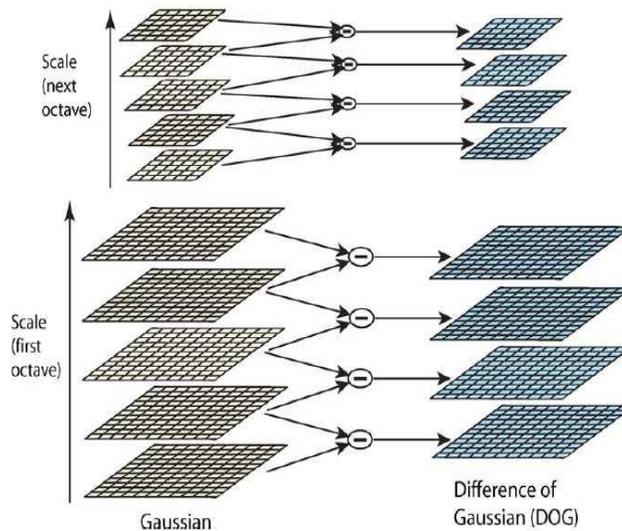


Figura 6 Proceso de cálculo de diferencia de gaussianas en diferentes escalas

El método SIFT consta de las siguientes etapas:

Construcción de pirámides de Scale-Space. El principal atractivo de la parametrización SIFT es la reducción de la imagen a un conjunto de parámetros invariantes a la escala y rotación. Esto se logra con la creación de pirámides en las que se varía la escala de la fotografía y con la convolución con una función gaussiana.

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma) \quad (1)$$

En la Figura 6 se muestra el proceso como se calculan las diferencias de gaussianas en las pirámides. La imagen I se convoluciona con una gaussiana G para obtener una imagen L un tanto más borrosa (1), luego se hace variar el parámetro S de la gaussiana y con esto se obtiene otra imagen aún más difusa.

La imagen obtenida por diferencia de gaussianas logra resaltar los bordes y esquinas de la imagen, como se aprecia en la Figura 7. Como las imágenes disponibles son de gran tamaño (1920 x 2560), la ilustración se calculó con una gaussiana con un valor S alto y una matriz cuadrada de convolución de orden 50.

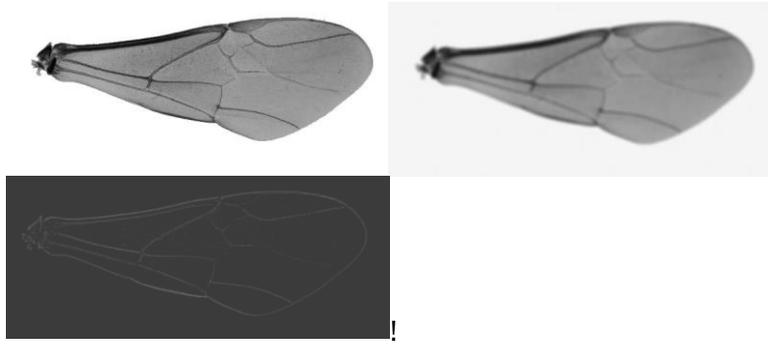


Figura 7 Ilustración del efecto del filtrado gaussiano y diferencia de gaussianas

Localización de puntos clave. Cuando ya se han calculado las diferencias gaussianas, se deben identificar los candidatos a máximos y mínimos entre los puntos que no se anularon. Es importante recalcar que la comparación de cada punto con sus vecinos se hace también en la escala próxima (Figura 8).

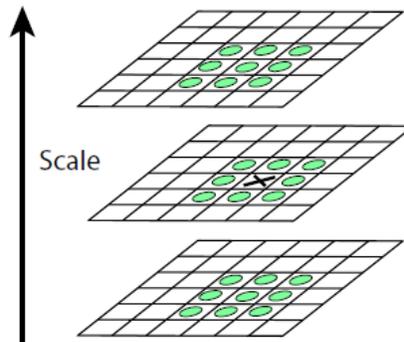


Figura 8 Comparación de puntos con sus vecinos para discriminar valores mínimos y máximos.

Estos puntos identificados como los máximos y mínimos de sus vecinos son luego discriminados según su estabilidad y clasificados como extremos a partir de una estimación D de Taylor para curvaturas cuadráticas en las diferencias gaussianas, el Hessiano, autovalores y un radio r que los relaciona (2).

$$\begin{aligned}
 H &= \begin{bmatrix} D_{xx} & D_{xy} \\ D_{yx} & D_{yy} \end{bmatrix} \\
 Tr(H) &= D_{xx} + D_{yy} = \alpha + \beta \\
 Det(H) &= D_{xx}D_{yy} - (D_{xy})^2 = \alpha\beta
 \end{aligned} \tag{2}$$

$$\frac{Tr(H)^2}{Det(H)} < \frac{(1+r)^2}{r}$$

En (2) α es el mayor de los autovalores de H y β es el menor de estos. Además, se tiene que $\alpha = r\beta$, con estas herramientas se identifican los puntos considerados claves con estabilidad y se descartan los ubicados a los largo de los bordes.

Asignación de orientación. Para asignar la orientación a los puntos (x, y) denominados claves en la etapa anterior, se utilizan las ecuaciones (3), en las que $m(x, y)$ es la magnitud del gradiente en (x, y) y $\theta(x, y)$ es la orientación.

$$m(x, y) = \sqrt{(L(x, y + 1) - L(x, y - 1))^2 + (L(x + 1, y) - L(x - 1, y))^2} \quad (3)$$

$$\theta(x, y) = \text{Tan}^{-1} \left(\frac{L(x, y + 1) - L(x, y - 1)}{L(x + 1, y) - L(x - 1, y)} \right)$$

Descriptor de puntos clave. Con la información obtenida en las etapas anteriores se asigna a cada punto clave un conjunto de características entre las que se destacan la posición, orientación y rotaciones del gradiente (para lograr invariancia respecto a rotaciones en 2D o cambios de puntos de vista en 3D). Cada punto tiene asociada una matriz de $4 \times 4 \times 8$, es decir 128 características (González & Woods, 2012), (Nixon, 2012).

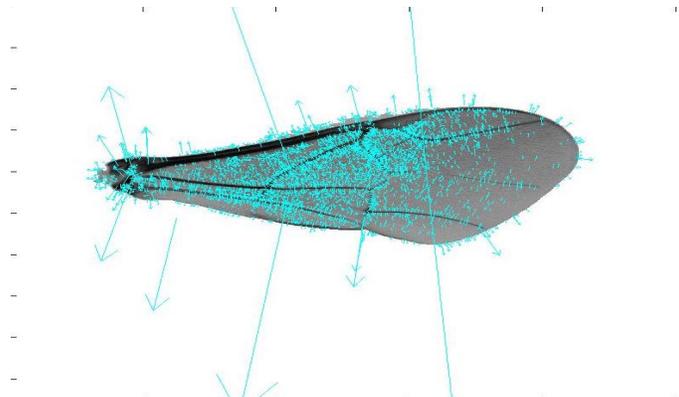


Figura 9 Salida de la parametrización SIFT para una foto con 4477 puntos claves.

Determinación de correspondencias (Match). Al aplicar las etapas anteriores a dos fotografías se tendrán una serie de puntos claves con su respectiva descripción en cada una de ellas. Lo siguiente es establecer si existe correspondencia entre las características de un punto clave con las características de alguno de la otra imagen, para esto se utiliza la distancia euclídea.

La idea propuesta por Lowe, consiste en determinar el vecino más próximo. Para un punto clave (x, y) en una imagen se calcula todas las distancias euclídeas desde los vectores directores de cada punto clave en la otra imagen hasta el vector director de (x, y) . Los casos

en los que las distancias no sobrepasen un umbral razonable serán consideradas correspondientes con (x, y) .

Es claro que esta estrategia no descarta la posibilidad de correspondencias múltiples, incorrectas y no correspondencia, sin embargo se ha considerado mejor tener suficientes correspondencias (aunque sean incorrectas) con las salidas SIFT (Figura 9) y luego afinar las asignaciones, por lo que se relajó un poco el umbral utilizado en el pareo.

7.2.3 Nuevo parámetro discriminador adicionado al método SIFT

El parámetro que usualmente se utiliza en el método SIFT para discriminar si un par de puntos son correspondientes, es la distancia promedio a todos los puntos claves generados. Sin embargo, este discriminador resultó insuficiente pues en muchos casos se hicieron asignaciones erróneas.

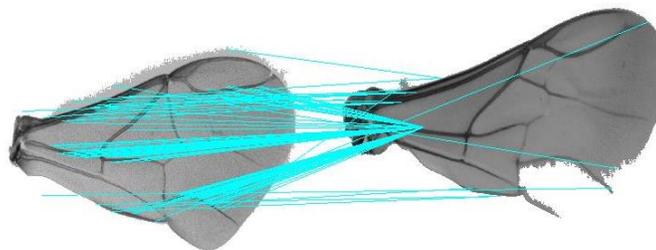


Figura 10 Caso extremo de correspondencias equivocadas entre las especies

Cephalotrigona zexmeniae y *Melipona beecheii*

Se tenía conciencia de que podrían suceder correspondencias equivocadas, aunque no se podía determinar en qué magnitud se darían. La Figura 10 muestra un caso donde el pareo resultó incorrecto para una gran cantidad de puntos claves. Aunque podría argumentarse que las fotografías no son adecuadas según los parámetros del protocolo de captura utilizado en el preprocesado, lo cierto es que este caso es caótico y sugiere la necesidad de incorporar algún discriminador más robusto para la etapa de pareo.

Es necesario, además, que el nuevo discriminador sea invariante al tamaño de la imagen y la rotación pues estas son condiciones establecidas por el método SIFT. Se decidió incorporar el siguiente discriminador de correspondencia un poco más exigente. Un par de puntos claves dados como correspondientes en la salida SIFT debían cumplir además con ser menores que un valor λ que se actualiza para cada asignación.

Para un punto clave con vector director (x_1, x_2) y otro punto clave (y_1, y_2) en otra imagen, considere

$$\begin{aligned}
 r_1 &= \sqrt{(x_1 - x_{1c})^2 + (y_1 - y_{1c})^2} \\
 r_2 &= \sqrt{(x_2 - x_{2c})^2 + (y_2 - y_{2c})^2} \\
 \rho_1 &= \frac{r_1}{A_1} \\
 \rho_2 &= \frac{r_2}{A_2} \\
 |\rho_1 - \rho_2| &< \lambda
 \end{aligned}
 \tag{4}$$

Donde (x_{1c}, y_{1c}) son las coordenadas del centro de masa de la imagen del ala y (x_1, y_1) es punto clave y análogamente para (x_{2c}, y_{2c}) . Las constantes A_1 y A_2 son las áreas de las correspondientes alas.

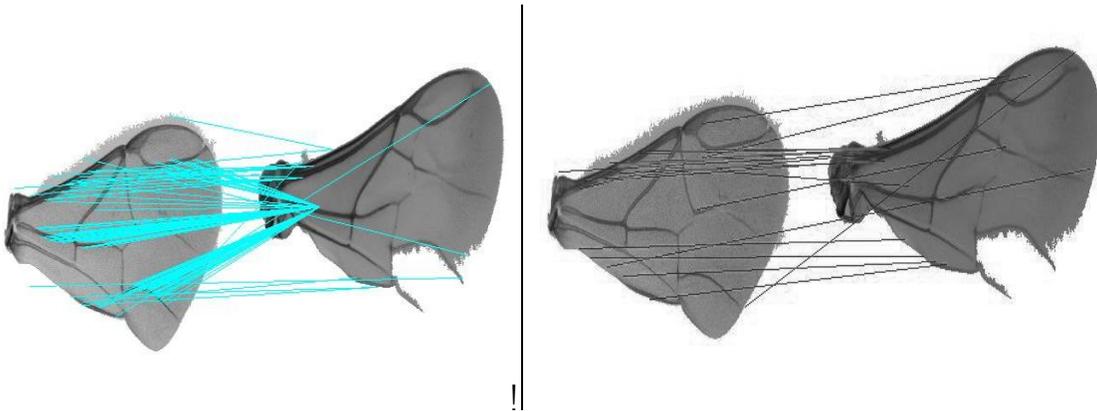


Figura 11 Efecto de la incorporación del nuevo discriminador en la correspondencias entre las especies *Cephalotrigona zexmeniae* y *Melipona beecheii*

El valor de λ es una cota a la diferencia entre los parámetros ρ descritos en (4). Los parámetros ρ son los radios normalizados medidos desde el vector director del punto clave hasta el centro de masa del ala. Para normalizar la distancia euclídea se utilizó un cociente por el área del ala. Con la incorporación de este discriminador se asegura la invariancia respecto de la escala y rotación y además se descartan asignaciones en las que los puntos tienen posiciones muy diferentes respecto de los centros de masa de las alas. La Figura 11 muestra la mejora en un en el caso extremo Figura 10, donde se habían dado muchas asignaciones incorrecta.

7.3 Desarrollo del sistema de clasificación

Partiendo de la base de datos con fotografías debidamente procesadas en la fase de eliminación del fondo y con la colaboración de biólogos especialistas en estudios apícolas, se seleccionó una fotografía por cada especie para crear una base de referencia para el sistema. Las fotos se seleccionaron con mucho cuidado considerando que fuesen una buena representación de la especie, desde el punto de vista biológico, y que las etapas previas de preprocesado han sido exitosas a la perfección (Herrera, y otros, 2013). Es decir, además de cumplir con los requerimientos de forma, color y texturas debe ser una foto en la que el ala se aprecie sin ruido y que la información que contenga sea únicamente del ala.

$$\begin{aligned}
 \text{Im} \Rightarrow [\text{Preprocesado}] \Rightarrow \text{Imp} \Rightarrow & \left. \begin{array}{l} I_1 \\ I_2 \\ \vdots \\ I_k \\ \vdots \\ I_n \end{array} \right\} \begin{array}{l} \text{Match}(\text{Imp}, I_1) \\ \text{Match}(\text{Imp}, I_2) \\ \vdots \\ \text{Match}(\text{Imp}, I_k) \\ \vdots \\ \text{Match}(\text{Imp}, I_n) \end{array} \\
 \\
 M_k = \max(\text{Match}) \Rightarrow & \left\{ \begin{array}{l} M_k \text{ es } \text{ú} \text{nico} \left\{ \begin{array}{l} M_k \geq N \Rightarrow \text{La abeja en la fotografía es de la especie } k \\ M_k < N \Rightarrow \text{Coincidencias insuficientes} \end{array} \right. \\ \\ M_k \text{ no es } \text{ú} \text{nico} \Rightarrow \text{Múltiples coincidencias} \end{array} \right.
 \end{aligned}$$

Figura 12 Esquema del sistema de discriminación de especies de abejas.

El sistema que se desarrolló (Figura 12) recibe una fotografía **Im** cualquiera que contenga un ala (o dos) de una abeja nativa costarricense (es importante que el ala pueda diferenciarse del fondo de la fotografía), la procesa para crear otra imagen **Imp** en la que el fondo se ha eliminado. La etapa siguiente es la aplicación de la parametrización SIFT a la nueva imagen **Imp**, para luego realizar un pareo con cada una de las fotos en la base de referencia. El sistema guarda el número de correspondencias obtenidas con cada foto **In** de la base, con esta información se calcula el máximo de estas correspondencias.

Si el máximo de correspondencia se dio con la imagen **Ik** entonces la foto que recibió el sistema se discrimina como la especie etiquetada en la imagen en la posición **k**. Es claro que podrían pasar algunas situaciones donde el sistema no pueda decidir; si el máximo de correspondencias aparece en varias imágenes o si solo está en una especie pero su valor es muy pequeño como para considerarse una discriminación exitosa. Estas situaciones están

previstas en el sistema, este reportaría de forma automática que hay ‘múltiples coincidencias’ o bien que se tienen ‘coincidencias insuficientes’. Para esta última situación se ha establecido un mínimo de coincidencias experimental N de 25 puntos claves.

El algoritmo fue programado en MatLab y se utilizó la librería vlfeat (disponible en <http://www.vlfeat.org/index.html>) para la aplicación del SIFT y las asignaciones de correspondencias.

8 Resultados

La estrategia para la obtención y análisis de resultados fue trabajar con una muestra pequeña de imágenes en las que se calibraron los procesos de eliminación de fondo, parametrización SIFT y determinación de correspondencias; luego se entrenó el algoritmo con una muestra más grande Tabla 2 y finalmente se validaron los cálculos con otra muestra de imágenes totalmente independiente a las utilizadas en las etapas anteriores.

Tabla 2 Totalidad de especies utilizadas en la etapa de entrenamiento.

Cant.	Especie	Cant.	Especie
8	<i>Cephalotrigona zexmeniae</i>	6	<i>Paratrigona ornaticeps</i>
5	<i>Dolichotrigona schulthessi</i>	8	<i>Scaptotrigona luteipennis</i>
8	<i>Geotrigona lutzi</i>	8	<i>Scaptotrigona pectoralis</i>
8	<i>Lestremielitta denuncia</i>	8	<i>Scaptotrigona subobscuripennis</i>
7	<i>Melipona beecheii</i>	9	<i>Tetragona zieglerei</i>
8	<i>Melipona costaricensis</i>	8	<i>Trigona corvina</i>
8	<i>Nannotrigona melleria</i>	5	<i>Trigona ferricauda</i>
8	<i>Nannotrigona perilampoides</i>	8	<i>Trigona fulviventris</i>
4	<i>Oxytrigona mellicolor</i>	8	<i>Trigona fuscipennis</i>
8	<i>Partamona orizabaensis</i>	8	<i>Trigona muzoensis</i>
8	<i>Plebeia frontalis</i>	8	<i>Trigona nigerrima</i>
8	<i>Plebeia jatiformis</i>	8	<i>Trigona silvestriana</i>

Los resultados de la etapa de entrenamiento se tabularon en una matriz de discriminación cruzada. Para la creación de esta matriz el algoritmo fue implementado en una muestra de 182 imágenes que incluían 25 especies.

Tabla 3. Porción de la matriz cruzada discriminadora del algoritmo en etapa de entrenamiento.

	Ce. ze 04	D. sc 07	G. lu 04	Geo. chi 02	L. da 07	Me. be 02	Me.co 03	Na. me 06	Na. pe 01	Ox. me 03	Par. or 01	Pl. fr 02	Pl. ja 10	Para. or 04	Scp. lu 05
Ce. ze 01	59	10	23	22	10	9	18	14	11	13	17	8	11	12	25
Ce. ze 02	52	11	17	24	9	11	20	11	6	10	10	6	6	18	32
Ce. ze 03	45	9	22	20	10	13	12	21	11	17	23	13	10	10	20
Ce. ze 04	22017	11	16	22	5	7	16	9	12	9	16	12	8	9	21
Ce. ze 05	57	18	16	38	15	12	32	16	7	36	30	13	14	11	59
Ce. ze 06	61	14	24	17	9	12	15	12	8	13	20	6	3	7	18
Ce. ze 07	66	4	16	13	6	4	20	5	11	12	20	10	6	10	17
Ce. ze 08	53	14	15	17	11	8	27	9	16	18	20	4	6	12	19
D. sc 07	38	3162	22	32	16	22	27	12	12	21	17	15	16	34	69
D. sc 22	11	19	9	27	4	12	14	2	11	8	5	4	5	15	14
D. sc 23	10	20	18	19	4	5	29	4	15	11	9	0	5	20	25
D. sc 24	0	2	0	13	0	0	0	0	0	0	0	0	0	11	0
D. sc 25	8	20	15	16	1	5	9	0	3	5	7	4	4	12	10
G. lu 01	21	7	41	32	9	7	21	19	12	18	14	11	9	7	22
G. lu 02	20	10	50	22	17	10	22	14	14	10	12	12	17	9	22
G. lu 03	21	9	42	27	1	11	19	19	7	11	11	16	6	9	21
G. lu 04	14	6	20771	23	7	10	17	13	8	10	15	6	8	13	12
G. lu 05	20	9	58	28	15	6	14	13	10	17	20	12	16	7	25
G. lu 06	16	7	46	29	9	4	23	16	8	16	16	8	10	15	20
G. lu 07	18	11	31	22	10	6	14	12	7	16	12	13	4	9	23
G. lu 08	17	16	46	33	6	8	19	16	13	21	12	12	8	13	27
Geo. chi 01	18	7	26	46	7	7	18	10	11	15	16	14	14	10	16
Geo. chi 02	21	11	19	15993	11	9	20	20	25	19	18	7	5	12	24
L. da 05	10	19	13	17	33	14	35	21	9	9	16	14	10	17	15
L. da 06	38	13	21	29	62	22	34	24	24	17	22	17	21	20	32
L. da 07	40	11	18	29	2193	18	49	26	28	29	36	32	17	26	22

En la Tabla 3 muestra una porción de la matriz cruzada de discriminación. En las filas se colocan las etiquetas que identifican las imágenes presentes en la muestra y en la columnas están las imágenes de referencia. El algoritmo calculó para cada imagen todas las coincidencias con las imágenes de referencia.

Para dar claridad al proceso se destacaron con color las celdas en donde debía obtenerse el máximo de correspondencias y los máximos se destacaron con color rojo. Si el número rojo está en una posición que no corresponde con el color destacado se considera un resultado equivocado. Por ejemplo, todas las imágenes de la especie *Geotrigona lutzi* (G. lu) fueron discriminadas perfectamente mientras que 3 de las imágenes de la especie *Dolichotrigona schulthessi* (D. sc) fueron asignadas a una especie equivocada.

La implementación del algoritmo en la etapa de entrenamiento tuvo un porcentaje de éxito del 87,4. Este nivel de éxito se ha considerado bueno para esta etapa, el algoritmo se acondicionó, se ajustaron los parámetros y se inició la etapa de validación en una muestra independiente de fotografías.

Para la validación de los resultados se utilizó una muestra de 459 fotografías que agrupan 18 especies en 10 géneros. Es importante aclarar que el tiempo destinado para este proyecto no permitió coleccionar una cantidad suficiente de especímenes de las 7 especies restantes como para cubrir las etapas de entrenamiento y validación. Por esta razón, aunque el algoritmo se entrenó con 25 especies, la etapa de validación solo puede incluir 18. Esto no es alarmante pues en las referencias consultadas ninguna ha llegado a esa cantidad de especies discriminadas.

Tabla 4 Porcentaje de éxito en la discriminación de abejas nativas costarricenses por género y por especie.

	Género	Especie	Codigo	Muestra	Errores	Ex.Espec	det. Genero
1	<i>Cephalotrigona</i>	<i>zexmeiae</i>	Ce.ze	11	0	100	100
2	<i>Geotrigona</i>	<i>lutzi</i>	Ge.lu	13	0	100	100
3	<i>Lestrimelitta</i>	<i>danuncia</i>	L.da	33	0	100	100
4	<i>Melipona</i>	<i>costaricensis</i>	Me.co	16	0	100	100
5	<i>Nanotrigona</i>	<i>mellaria</i>	Na.me	35	2	94,2857	95,1219512
6		<i>perilampoides</i>	Na.pe	6	0	100	
7	<i>Partamona</i>	<i>orizabaensis</i>	Par.or	32	1	96,875	96,875
8	<i>Plebeia</i>	<i>frontalis</i>	Pl.fr	17	13	23,5294	100
9		<i>jatiformis</i>	Pl.ja	27		100	
10	<i>Scaptotrigona</i>	<i>pectoralis</i>	Scp.pe	37		100	98,5294118
11		<i>subobscuripennis</i>	Scp.su	31	1	96,7742	
12	<i>Tetragona</i>	<i>ziegleri</i>	Te.zi	24	3	87,5	87,5
13	<i>Trigona</i>	<i>corvina</i>	Tri.co	30	6	80	97,1751412
14		<i>fulviventris</i>	Tri.ful	36	14	61,1111	
15		<i>fuscipennis</i>	Tri.fus	32	3	90,625	
16		<i>muzoensis</i>	Tri.mu	33		100	
17		<i>nigerrima</i>	Tri.ni	16	10	37,5	
18		<i>silvestriana</i>	Tri.si	30	9	70	
				459	62	86,4924	97,5201504

El porcentaje de éxito en la discriminación de especie resultó en 86,5 mientras que la discriminación por género llegó al 97,52% (Tabla 4). Con un análisis detallado se logró constatar que las asignaciones de especies equivocadas se hicieron en su mayoría en el mismo género, lo que lleva a considerar que las alas son realmente muy parecidas en algunas especies, lo cual sugiere que se debe tomar en cuenta otros parámetros para su clasificación

9 Discusión y conclusiones

En un sistema de discriminación se espera que el porcentaje de éxito sea cercano a 100. Con la discriminación de abejas nativas sin aguijón a nivel de género se ha cumplido con esta expectativa, sin embargo, es necesario mejorar el porcentaje de identificación de especie.

Con la idea de mejorar el porcentaje de éxito para la discriminación de especies se consultó con los biólogos del CINAT, expertos en estudios apícolas, y se decidió que es necesario incorporar algunos discriminadores de especies que no están en el ala, tales como colores de otras partes del cuerpo, vello, tamaño de patas o antenas u otras. Al igual que en otras abejas sin aguijón, como *Paratrigona Schwarz* y *Plebeia sp.*, especies del género *Nogueirapis* se reconocen principalmente por las diferencias en el patrón de color de las obreras. Tales diferencias de color son a menudo apoyados por otros rasgos morfológicos incluidos el tamaño del cuerpo, así como la forma y proporciones de las diferentes estructuras membranosas. La propuesta es hacer que el sistema cuando tenga 'duda' sobre la especie, ya sea por la existencia de valores muy cercanos al máximo de coincidencias o por coincidencias múltiples, se torne en un sistema semi asistido. Por ejemplo, se ha determinado que las especies *Trigona silvestriana* y *Trigona fulviventris* tienen alas muy similares que podrían confundir al sistema, sin embargo estas especies se diferencian en la coloración del abdomen.

Con algunas preguntas sencillas al usuario, sin necesidad de entrenamiento ni gasto excesivo en laboratorios, el sistema mejoraría sin duda el porcentaje de detección por especie. Otra posibilidad interesante es considerar la georreferencia de la imagen y que esta ingrese al sistema como un nuevo parámetro discriminador.

Además, se ha pensado que a futuros el sistema podría ser llevado a ambientes menos controlados y ponerlo a disposición de los meliponicultores. Sería un gran avance la implementación del sistema en algún dispositivo móvil y que el preprocesado identifique el ala de la abeja incluso con la presencia del resto del cuerpo. Además, con el diseño de una interfaz para el usuario y el mantenimiento en un servidor los meliponicultores e investigadores apícolas tendrían una herramienta muy valiosa que podría ser usada incluso vía web.

10 Referencias

- Arbuckle1, T., Schröder, S., Steinhage, V., & Wittmann, D. (2001). Identification and Monitoring of bee Species using ABIS. *Proceedings of the 15th International Symposium Informatics for Environmental Protection* (págs. 425-430). Zurich: EMPA.
- Ayala, R., & Engel, M. (2014). A new stingless bee species of the genus *Nogueirapis* from Costa Rica (Hymenoptera: Apidae). *Jouenal of Melittology*, 1-9.
- Blacquièrre, T., Smaghe, G., van Gestel, C., & Mommaerts, V. (2012). Neonicotinoids in bees: a review on concentrations, side-effects and risk assessment. *Ecotoxicology*, 973-992.
- Breuker, C., Gibbs, M., Van Dongen, S., Merckx, T., & Van Dyck, H. (2010). The Use of Geometric Morphometrics in Studying Butterfly Wings in an Evolutionary Ecological Context. En A. Elewa, *Morphometrics for Nonmorphometricians* (págs. 271-287). Berlin: Springer.
- Brown, M., & Paxton, R. (2009). The conservation of bees: a global perspective. *Apidologie*, 410-416.
- Camargo, J., & Pedro, S. (2007). Meliponini Lepeletier. En J. Moure, & D. Urban, *Catalogue of bees (Hymenoptera, Apoidea) in the neotropical region* (págs. 272-578). Curitiba: Sociedade Brasileira de Entomologia.
- Francoy, T., Silva, R., Nunes-Silva, P., Menezes, C., & Imperatriz-Fonseca, V. (2007). Gender identification of five genera of stingless bees (Apidae, Meliponini) based on wing morphology. *Genetics and Molecular Research*, 207-214.

- Francoy, T., Wittman, D., Muller, S., Steinhage, V., Bezerra-Laure, V., De Jong, D., & Goncalves, L. (2008). Identification of africanized honey bees through wing morphometrics: two fast and efficient procedures. *Apidologie*, 488-494.
- Freitas, B., Fonseca, V., Medina, L., Peixoto K., A., Galetto, L., Parra, J., . . . Quesada Euán, J. (2009). Diversity, threats and conservation of native bees in the neotropics. *Apidologie*, 332-346.
- Gonzalez, J. (2000). *Visión por computadora*. Madrid: Paraninfo.
- González, R., & Woods, R. (2012). *Digital Image Processing using MatLab*. New York: Prentice Hall.
- Hall, C. (2011). *Automated approach to bee identification from wing venation*. Madison: Universidad de Wisconsin.
- Hall, M., MacLeod, N., & Wardhana, A. (2014). Use of wing morphometrics to identify populations of the Old World screwworm fly. *Elsevier*, 549-555.
- Herrera, E., Aguilar, I., Prendas, P., Figueroa, G., Travieso, C., & Ramírez, M. (2013). Photograph capture protocol for the creation of an automatic classification. *VIII Congreso Mesoamericano de Abejas Nativas* (pág. 337). Heredia: Universidad Nacional de Costa Rica.
- Meléndez R., V., Meneses C., L., & Kevan G., P. (2013). Effects of human disturbance and habitat fragmentation on stingless bees. En P. Vit, P. Silvia, & D. Roubik, *Pot-Honey: A legacy of stingless bees*. New York: Springer.
- Michener D., C. (2007). *The bees of the world*. New York: Johns Hopkins University Press.
- Nixon, M. (2012). *Feature Extraction and Image Processing for Computer Vision*. New York: Academic Press.
- Prendas, P., & Ramírez, M. (2011). *Algoritmo para la detección y seguimiento del ácaro v destructor en secuencia digital de video*. Heredia: Universidad Nacional de Costa Rica.
- Ramírez, P., & Bernal, A. (2005). Protocolo para la preservación y manejo de colecciones biológicas. *Boletín científico centro de museos*, 117-148.
- Thompson, H. (2012). Behavioural Effects of Pesticides in Bees—Their Potential for Use in Risk Assessment. *Ecotoxicology*, 317-330.
- Tofilski, A. (2008). Automatic measurements of honeybee wings. En N. MacLeod, *Automated Taxon Identification in Systematics: Theory, Approaches and Applications* (págs. 289-298). Florida: CRC Press.
- Vaissière, B., Klein, A., Cane, J., & Steffan-Dewenter, I. (2006). Importance of pollinators in changing landscapes for word crops. *Proceedings The Royal Society B*, 303-313.

Whitehorn, P., O'Connor, S., Walkers, F., & Goulson, D. (2012). Neonicotinoid pesticide reduces bumble bee colony growth and queen production. *Science* 336, 351-352.

Yüksel, S., & Tüzün, A. (2011). Use of wing morphometry for the discrimination of some *Cerceris* (Insecta: Hymenoptera: Crabronidae species). *African Journal of Biotechnology*, 18569-18580.