

MÉTODOS Y ALGORITMOS DE PROCESAMIENTO DE IMÁGENES QUIMIOMÉTRICAS Y BIOMÉTRICAS.

UNIDAD EJECUTORA PRINCIPAL DEL RESULTADO: Centro de Aplicaciones de Tecnologías de Avanzada (División de Investigaciones, CENATAV, Datys Soluciones Tecnológicas)

AUTOR PRINCIPAL: Dr.C. Francisco José Silva Mata (CENATAV, Datys Soluciones Tecnológicas)

OTROS AUTORES: Dr.C. Isneri Talavera Bustamante(CENATAV, Datys Soluciones Tecnológicas) Dr.C. Ángel Augier Calderín (Universidad de la Habana)

COLABORADORES: Dr.C. Stefano Berretti, Universidad de Florencia, Italia. Dr.C. Alberto del Bimbo, Universidad de Florencia, Italia. Dr.C. Pietro Pala. Universidad de Florencia. Italia. Dra.C. Heydi Méndez Vázquez, CENATAV. Dr.C. Jose Luis Gil, CENATAV. Dr. C. Eduardo Garea Llano CENATAV. Dr. C. Yenisel Plasencia Calaña CENATAV. Dr. C. Edel García CENATAV. Ing. Yoanna Martinez Diaz CENATAV. Ing. Ricardo González Gazapo CENATAV. Ing. Dania Porro Muñoz CENATAV. Dr. C. Diana Porro Muñoz CENATAV. Lic. Estela María Alvarez Morales UH. Dr.C. Noslen Hernández González CENATAV. Lic. Anier Revilla Eng CENATAV. Lic. Victor Manuel Mendiola Lau CENATAV. Lic. Juan Ramiro Palau CIDT. Lic. Maria Santiesteban Vidal. Dr.C. Marcia Ferreira González UNICAMP, Brasil. Dr. C. Raul Alvarez Valdés FAR. Ing. Moisés Herrera ULV. MsC. René Antonio Vázquez Vallejo LCC. Lic. Yenisel Pérez Pérez LCC

AUTOR PARA LA CORRESPONDENCIA:

Dr. Francisco José Silva Mata

Dirección: 7ma #21406 e/ 214 y 216, Reparto Siboney, Playa, C.P. 12200

Fax: (+) 537 273 0045

Correo electrónico: fjsilva@cenatav.co.cu

RESUMEN

En el área de reconocimiento de patrones, se reporta la creación y desarrollo de nuevos métodos y algoritmos de procesamiento de imágenes quimio-métricas y biométricas que permiten el incremento, con respecto a estándares internacionales, de la eficiencia y la eficacia del reconocimiento de sustancias y personas respectivamente. Estos métodos han sido integrados en las secuencias de análisis específicas para la detección, normalización, segmentación, representación y clasificación, en la automatización del procesamiento de imágenes de geles de poliacrilamida para la extracción de perfiles de ADN, en el análisis de placas de geles, y en la identificación de personas mediante el uso de imágenes de rostro e iris.

Las características quimio-métricas se han representado mediante vectores que describen las características de las manchas en la imagen. Para las características biométricas se crearon dos nuevas representaciones, una basada en distancias geodésicas, que permite cotejar directamente imágenes bidimensionales de rostros con sus análogas tridimensionales, y otra basada en clases de funciones, que mostró gran eficacia en el reconocimiento de rostros tridimensionales y de iris, reduciéndose notablemente la dimensión de los datos, y obteniendo alta eficiencia de los procesos.

Parte de estos resultados han sido introducidos en la práctica como parte de las aplicaciones de identificación de perfiles de ADN a partir de imágenes de poliacrilamida, en las aplicaciones de reconocimiento de drogas producidas por la empresa DATYS y en general en aplicaciones de procesamiento de imágenes formando parte de los productos de esta empresa.

Los resultados teóricos obtenidos han sido presentados en seis artículos en revistas internacionales de impacto (referenciados en scopus), dos capítulos de libros en las editoras IGI Global y Springer, en un total de 12 congresos internacionales de impacto y publicados en sus memorias (IEEE, LNCS), además parte de estos trabajos fueron presentados en tres congresos internacionales y cinco nacionales celebrados en Cuba. La novedad científica del trabajo está avalada porque estos resultados están contenidos en la Tesis de Doctor en Ciencias Técnicas del autor principal.

COMUNICACIÓN CORTA DEL RESULTADO

El impacto de estas tecnologías está cambiando la filosofía de la solución de los problemas de otras ciencias o de las ramas de estas en diversos campos de la actividad humana. Entre las ramas donde la incidencia del desarrollo de las tecnologías computacionales ha sido decisivo, pueden mencionarse aquellas relacionadas con la identificación de sustancias químicas o su análisis, conocida como Quimiometría y la identificación de personas, conocida como Biometría.

La quimiometría se encarga de los procesos de identificación y análisis de sustancias, algunos de los cuales están basados en placas, que son comúnmente analizadas de forma manual. A tono con el desarrollo de las nuevas tecnologías, estas placas pueden ser fotografiadas y posteriormente procesadas automáticamente, mediante sistemas que básicamente realizan funciones comunes a la generalidad de los sistemas de procesamiento de imágenes cuyos fines son identificativos. En las imágenes de placas se presentan problemas como la presencia de sombras o de una iluminación inadecuada, los errores provocados por la mala manipulación de los reactivos y las variaciones propias de los procesos físico químicos, los bajos contrastes de las manchas respecto al fondo causado por la fluctuación de la concentración de las sustancias, los cambios de coloración, la deformación geométrica de los carriles, etc.

Análogamente los sistemas biométricos tienen como objetivo la determinación o comprobación de la identidad de personas por sus características o por su comportamiento. Debido a su poder identificativo, la posibilidad de su registro de forma no invasiva y sus posibilidades de ser utilizado en aplicaciones públicas o encubiertas, el rostro humano se perfila como una característica biométrica de amplia utilización. Una

de sus partes, el iris del ojo, está siendo utilizado también para la identificación automática de personas. Sin embargo, existen problemas que dificultan el buen desempeño de los sistemas mencionados. En el caso del rostro existen inconvenientes típicos de los sistemas de reconocimiento, los cuales están relacionados con la gran variabilidad de la apariencia facial, debido a las poses, el envejecimiento, las fuentes de iluminación y los cambios de expresiones faciales. Estos elementos producen una alta variabilidad intra-sujeto, una gran complejidad y no convexidad de las variedades (en inglés, manifolds) y una alta dimensión, la cual resulta contrastante con el pequeño número de imágenes de rostro disponibles de la misma persona, lo que hace difícil el aprendizaje y generalización de la variedad de estos. Igualmente en el caso del iris se producen estos problemas, que afectan el desempeño de los sistemas automáticos para su reconocimiento.

En las aplicaciones automáticas de procesamiento de imágenes, tanto quimiométricas como biométricas, las entradas son imágenes y sus esquemas de trabajo son representados mediante etapas lógicas de funcionamiento, las cuales conforman una secuencia completa de procesamiento de la imagen original, desde su captura, hasta la identificación de la característica quimiométrica o biométrica.

La identificación de la característica quimiométrica significa en este contexto, responder a qué persona corresponde el ADN o qué sustancia patrón previamente almacenada coincide con la muestra. La identificación de la característica biométrica significa responder a qué persona pertenece la característica analizada respecto a las registradas previamente en una base de datos o fichero.

El trabajo preciso, rápido y bien correlacionado de todas las etapas, condiciona la calidad en la respuesta final de estos procesos. Resulta entonces imprescindible resolver las dificultades previamente mencionadas, para poder contar con sistemas de reconocimiento que resulten competitivos y utilizables en el entorno actual del desarrollo emergente de las tecnologías de la información y las comunicaciones, logrando a su vez una alta eficiencia medida a través de la complejidad temporal o tiempo de ejecución del proceso y de la complejidad espacial o espacio de memoria requerido y una alta eficacia de la respuesta (medida a través de las tasas de reconocimiento de estos sistemas; sustituyendo siempre que sea posible, las operaciones manuales por operaciones automáticas.

Teniendo en cuenta los aspectos referidos, se determinó como problema de investigación que los sistemas de procesamiento de imágenes para el reconocimiento de sustancias (basados en geles) y de personas (basados en el rostro y el iris), no alcanzan en muchos casos tasas de reconocimiento y rapidez adecuadas, como resultado de la insuficiencia de los métodos en cada una de las etapas de los procesos o de su escasa automatización.

Para dar solución a este problema se planteó como objetivo general de esta investigación la creación y desarrollo de nuevos métodos y algoritmos de procesamiento de imágenes quimiométricas y biométricas destinados a lograr incrementar la eficacia y la eficiencia de los procesos correspondientes de análisis e identificación de sustancias y de personas. Se ha centrado la atención específicamente en los procesos quimiométricos de identificación de perfiles de ADN sobre geles de poliacrilamida, en la identificación de sustancias mediante cromatografía de capa delgada y en los procesos de reconocimiento biométrico de rostro e iris.

NOVEDAD CIENTÍFICA

La actualidad y novedad científica del tema radica en la creación y desarrollo de nuevos métodos y algoritmos, que aplicados en cada una de las etapas del procesamiento de las imágenes e integrados apropiadamente, logran mejoras sustanciales aplicables en el desarrollo de sistemas de identificación quimiométrica y biométrica. Sobre la base de los problemas detectados en la literatura relacionada, en el marco de la investigación, se obtuvieron los siguientes resultados los cuales constituyen aportes en esta área de investigación:

1. Nuevo método de extracción automática de perfiles de ADN a partir de imágenes de geles de poliacrilamida.
2. Nuevo algoritmo de restauración de manchas patrones de regiones "STR loci".
3. Nueva representación vectorial de manchas de ADN basada en descriptores de regiones y de contorno.
4. Nuevo método de separación de manchas solapadas, mediante la detección de cadenas de Freeman típicas.
5. Nuevo método de determinación de subregiones "STR loci" mediante el algoritmo de agrupamiento líder secuencial".
6. Nuevo método de segmentación de manchas basado en la variación de la pendiente de la envolvente del histograma y en el contexto.
7. Nueva método de descripción de manchas sobre geles.
8. Nuevo método de normalización fotométrica de imágenes de rostro afectadas por la luz especular, que combina la detección rápida de regiones afectadas, y la modificación de los valores de intensidad interiores de la región
9. Nuevas expresiones para la evaluación de la precisión del proceso de detección de objetos y regiones en imágenes.
10. Nuevo método de representación de imágenes de rostros 2D y modelos 3D que permite el cotejo directo entre estos para el reconocimiento.
11. Nuevo método de normalización de rostros para el cotejo de modelos 2D-3D.
12. Nuevo método de reconocimiento de rostro 3D mediante análisis de datos funcionales.
13. Nuevas cascadas de detectores para la detección del borde exterior del iris y la pupila, directamente integrables a las aplicaciones basadas en OpenCV.
14. Nuevo método de segmentación rápida de la zona codificable del iris.
15. Nueva secuencia de reconocimiento del iris mediante la aplicación de análisis de datos funcionales basada en dos representaciones de su textura utilizando como base los polinomios de Zernike anulares y circulares.
16. Nueva plataforma para el reconocimiento de iris.

IMPACTO DEL RESULTADO

El desarrollo de secuencias de procesamiento de imágenes de geles permite automatizar el procesamiento de los datos resultantes de los procesos experimentales de investigación de la identidad de sustancias y la identificación de perfiles de ADN, sustituyendo la intervención manual del hombre, reduciendo los errores subjetivos y aumentando la velocidad de las respuestas. Los aportes en esta dirección tienen un impacto importante, ya que resuelven algunos de los problemas que presentan las técnicas químico-analíticas basadas en placas, como el ruido, la falta de homogeneidad, las variaciones propias de los procesos físico-químicos y sus mediciones; los cuales perjudican sus desempeños, automatizando casi totalmente los procesos identificativos. La investigación realizada, presenta avances a nivel internacional en la forma de nuevos métodos y algoritmos propios y competitivos, comparados con los reportados en el estado del arte, que superan en algunos casos los indicadores de eficacia y eficiencia publicados recientemente, al resolver parcial o totalmente los problemas mencionados. Los resultados teóricos alcanzados constituyen una contribución científica al desarrollo de la biometría y la quimiometría, aportando referentes a nivel internacional y formando parte algunos de ellos de soluciones tecnológicas de la empresa DATYS.

La importancia que reviste esta investigación a partir de su introducción en la práctica, radica en el aporte de soluciones propias, que permiten su aplicación en los sistemas de identificación, videovigilancia y análisis forense, entre otros, dotando al país con nuevas herramientas computacionales para las aplicaciones que se suman al esfuerzo por obtener una independencia tecnológica, en una rama que reviste importancia especial para el enfrentamiento al delito y el mantenimiento de la seguridad y el orden interior. Los aportes introducidos se encuentran publicados en eventos y revistas de alto impacto internacional, lo que contribuye a la visibilidad de la labor científica de nuestro país.

PUBLICACIONES MÁS IMPORTANTES EN EL MARCO DE LA INVESTIGACIÓN

¹ F. J. Silva-Mata, Dania P. Munoz, V. M. Mendiola-Lau, I. Talavera-Bustamante, A.

G. Augier. "Criterios y métodos de selección de bases y su impacto en el análisis de datos funcionales. Algunos ejemplos en Biometría". Revista Cubana de Física, Volumen 33, No. 1e, 2016. (Scopus)

² F. J. Silva-Mata, Dania P. Munoz, S. Berretti, V. M. Mendiola-Lau, I. Talavera-Bustamante, N. Hernández-González, Y. Martánez-Díaz, A. G. Augier. "Alineación de senales e imágenes durante la aplicación del análisis de datos funcionales. Ejemplos prácticos de senales e imágenes quimiométricas y biométricas". Revista Cubana de Física, Volumen 33, No. 1e, 2016. (Scopus)

³ Dania P. Munoz, F.J. Silva-Mata, N. Hernández-González, I. Talavera-Bustamante. "Functional Data Analysis as an Alternative for the

Automatic Biometric Image Recognition: Iris Application". Rev. Computación y Sistemas, CONACYT, México, Vol. 18 No. 1, pp. 111-121, ISSN 1405-5546, 2014 (Scopus)

4 F.J. Silva-Mata, D. Porro-Munoz, D. Porro-Munoz, N. Hernández, I. Talavera-Bustamante, Y. Martínez-Díaz, L. Bustio-Munoz. "Integrating Computational Methods for Forensic Identification of Drugs by TLC, GC and UV Techniques". A.K. Muda et al. (eds.), Computational Intelligence in Digital Forensics: 187 Forensic Investigation and Applications, Studies in Computational Intelligence 555, Springer International Publishing Switzerland, pp. 187-210, 2014. (Scopus)

5. F.J. Silva-Mata, D. Porro-Munoz, I. Talavera-Bustamante, N. Hernández, R. Vázquez-Vallejo, Y. Pérez., "El procesamiento de imágenes de TLC como una vía para la identificación de sustancias". Revista Cubana de Química Vol XXV N. 3. 2013.(Revista certificada por el Citma) <http://www.redciencia.cu/>

6. S. Berretti, A. Del Bimbo, P. Pala, F.J. Silva-Mata, "Face Recognition based on Manifold Learning and SVM Classification of 2D and 3D Geodesic Curves". in Machine Learning Techniques for Adaptive Multimedia Retrieval: Technologies Applications and Perspectives, chap. 4, pp. 62-81, IGI Global, 2011.

7. Eduardo Garea Llano, F. J. Silva-Mata, N. Hernández-Gonzalez, R. González-Gazapo, J. Palau and Marcia M. Castro-Ferreira. "Application of chemometric tools for automatic classification and profile extraction of DNA samples in forensic tasks". Revista Analytica Chimica Acta, Vol. 595, Issues 1-2, pp. 43-50, 9 July 2007.

8 I. Talavera-Bustamante, F. J. Silva-Mata Eduardo Garea Llano, N. Hernández-González, R. González-Gazapo. "Improvement the oxyribo nucleic acid spots classification in polyacrilamide gel images using photometric normalization algorithms". Revista Analytica Chimica Acta, Vol. 595, Issues 1-2, pp. 145-151, 9 July 2007.

9 Dania P.Munoz, F.J. Silva-Mata, N. Hernández-González, I. Talavera-Bustamante. "3D Face Recognition by Functional Data Analysis". In proc. of CIARP 2014, , Puerto Vallarta, México, LNCS 8827, pp. 818–826, 2014.

10 Dania P.Munoz, F.J. Silva-Mata, N. Hernández-González, I. Talavera-Bustamante. "A New Iris Recognition Approach Based on a Functional Representation". In Proc. of CIARP 2013, Habana, Cuba, Part II, LNCS 8259, pp. 391–398, 2013.