

Detección y Segmentación de Franjas en Proyecciones de Luz Estructurada por Medio de Visión por Computadora

Juan Pablo Guerra Ibarra¹

Ana Celia Segundo Sevilla²

Alicia Linares Ramírez³

¹Instituto Tecnológico de Estudios Superiores de Zamora, jp@teczamora.mx

²Instituto Tecnológico de Estudios Superiores de Zamora, asegundo@accitesz.com

³Instituto Tecnológico de Estudios Superiores de Zamora, alinares@accitesz.com

Detección y Segmentación de Franjas en Proyecciones de Luz Estructurada por Medio de Visión por Computadora

Resumen

En este trabajo se abordan temáticas referentes a Visión por computadora o visión artificial, que son aplicadas a imágenes con proyección de franjas de luz estructurada, con la finalidad de prepararlas para la recuperación de la forma del objeto que se encuentre bajo el haz de luz.

La visión por computadora es un campo de la inteligencia artificial, el cual consiste esencialmente en programar una computadora para que "entienda" una escena o las características de una imagen. Los objetivos de la visión artificial incluyen, entre otras acciones, la detección, segmentación, localización y reconocimiento de ciertos patrones dentro de una serie de imágenes [1,6].

El área de Visión por computadora en los últimos años ha tenido un desarrollo importante y tomado cada vez mayor relevancia en los desarrollos tecnológicos que impactan en la vida diaria. Para lograr lo anterior las imágenes deben ser preprocesadas para con ello poder extraer de esta información de diferentes tipos, como lo son: distancia entre objetos, velocidad, temperatura, forma de objetos, entre otros.

El procesamiento digital de imágenes es el conjunto de técnicas que se aplican a las imágenes digitales con el objetivo de mejorar la calidad o facilitar la búsqueda de información con la finalidad de estudiarla. El análisis consiste en la extracción de propiedades y características de las imágenes, así como la clasificación e identificación y el reconocimiento de patrones. [2].

Palabras clave: Luminosidad, Imágenes, Visión, Artificial, Virtuales.

Strip Detection and Segmentation in Structured Light Projections through Computer Vision

Abstract

This work deals with topics related to computer vision or artificial vision, which are applied to images with projection of structured light strips, in order to prepare them for the recovery of the shape of the object that is under the beam of light.

Computer vision is a field of artificial intelligence, which essentially consists of programming a computer to "understand" a scene or the characteristics of an image. The objectives of artificial vision include, among other actions, the detection, segmentation, localization and recognition of certain patterns within a series of images [1].

The area of Vision by computer in the last years has had an important development and taken more and more relevance in the technological developments that impact in the daily life. To achieve the above, the images must be preprocessed in order to be able to extract from these information different types, such as: distance between objects, speed, temperature, shape of objects, among others.

Digital image processing is the set of techniques that are applied to digital images in order to improve the quality or facilitate the search for information in order to study it. The analysis consists in the extraction of properties and characteristics of the images, as well as the classification and identification and the recognition of patterns. [2].

Keywords: Luminosity, Images, Vision, Artificial, Virtual.

Introducción

La Inteligencia artificial es una rama de estudio que busca emular las capacidades humanas para comprender su medio ambiente. Una de las temáticas con más potencial es la visión por computadora debido a la diversidad de aplicaciones que se pueden tener.

La visión por computadora tiene como finalidad la extracción de información por medio de la manipulación de imágenes digitales. Una imagen se puede definir como una función de intensidad de luz bidimensional $f(x,y)$, donde “x ” e “y ” indican las coordenadas espaciales y el valor de f en cualquier punto (x,y) es proporcional a la luminosidad de la imagen en dicho punto[3].

La proyección de luz estructurada es un método de interferometría óptica, en la cual se proyecta un patrón franjas de luz sobre un objeto el cual crea interferencia en el patrón de franjas y la medición de estas brinda información acerca del objeto. Estas interferencias son capturadas en una imagen por medio de una cámara digital, debido al proceso de captura se genera ruido debido a diferentes causas, lo que puede derivar en información incorrecta al momento de su extracción.

En el proceso de preparación de la imagen para la extracción de información se emplean diferentes algoritmos de Visión por computadora dependiendo del tipo de corrección que se desee llevar a cabo.

En este trabajo se abordan algunas técnicas de visión por computadora para la eliminación de ruido en imágenes con patrones de franjas para su posterior segmentación.

Objetivo General

Realizar un estudio e implementación de las técnicas y algoritmos de visión por computadora que permitan detectar y segmentar las franjas de un patrón de luz estructurada proyectada sobre diferentes objetos, con la finalidad de tener un conocimiento acerca del objeto analizado.

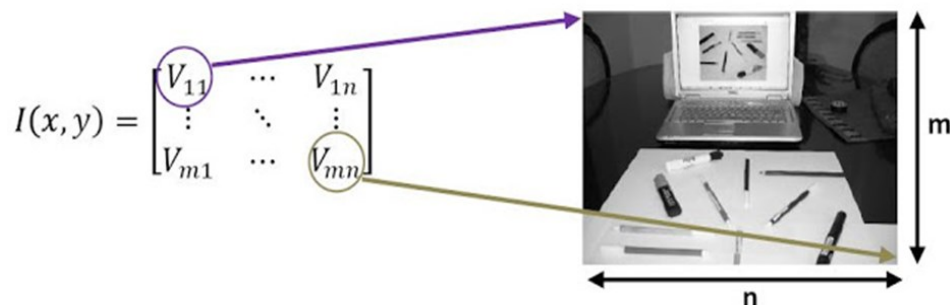
Marco teórico

Imágenes digitales

Una imagen se puede definir como una función de intensidad de luz bidimensional $f(x, y)$ o $I(x, y)$, donde “x” e “y” indican las coordenadas espaciales y el valor de en cualquier punto (x, y) es proporcional a la luminosidad (o nivel de gris) de la imagen en dicho punto. [1]

Una imagen digital se puede expresar como una función $f(x, y)$ o $I(x, y)$ que ha sido discretizada tanto en coordenadas espaciales como en luminosidad, como lo muestra la figura 1.1. Una imagen digital puede ser considerada como una matriz cuyos índices de renglón y columna identifican un punto (un lugar en el espacio bidimensional) en la imagen y el correspondiente valor de elemento de matriz determina el nivel de intensidad luminosa en aquel punto. Los elementos de estos arreglos digitales son llamados elementos de imagen o píxeles.

Figura 1: Representación de una imagen digital



Para las imágenes a color, la definición es prácticamente la misma, la única diferencia es que miden la intensidad a tres longitudes de onda distintas, llamadas colores, resultando tres valores de intensidad distintos para cada “pixel”, esto se interpreta como la intensidad de cada color en este punto, por lo que al sumar las tres intensidades se obtiene el color real del “pixel”, y de la misma forma se hace para el resto de la imagen. (En este proyecto se trabajó con imágenes en escala de grises, es decir, monocromáticas).

Imágenes de intensidad (Escala de grises)

Para el registro de una imagen se acostumbra a usar formatos que están referenciados por el número de bits que son utilizados para la cuantización de la cantidad de luz recibida. Cuando se trabaja solo la intensidad de la luz recibida en el detector se dice que la imagen ha sido adquirida en *tonos de gris*, los tipos más frecuentes se muestran en la tabla 1.1

Tabla 1: Tipos de bits y descripción

<i>Número de bits</i>	<i>Descripción</i>
1	Imagen Monocromática / Blanco y Negro / Alto contraste
4	Imagen de 16 niveles
8	Imagen de 256 niveles

Actualmente la mayor parte de los dispositivos simples permiten adquirir la imagen con una profundidad de 8 bits, los cuales permiten representar 256 niveles de gris, el rango es de $\{0,1,\dots,254,255\}$. Bajo esta representación los elementos de la matriz que representa a la imagen $f(x,y) \in \{0,1,\dots,254,255\}$ Al punto $x_{xy} = f(x,y)$ se le llama “pixel” y se debe entender el tono de gris que se ha asociado a la imagen en la coordenada (x,y) de la partición definida por el sistema de digitalización.

Figura 2: Imagen en escala de grises**Filtrado de orden**

Los filtros de orden están basados en un tratamiento específico de la estadística de la imagen llamado estadística de orden. Estos filtros operan en una vecindad de un determinado píxel, denominada ventana y reemplazan el valor del píxel central. La estadística de orden es una técnica que organiza todos los píxeles de la ventana en orden secuencial, basándose en el valor de nivel de gris de los mismos. El lugar de un determinado valor dentro de este conjunto está ordenado es referido como el rango. Dada una ventana V de dimensión $n \times n$, los valores de los píxeles pueden ordenarse de menor a mayor como siguen:

$$f_1 \leq f_2 \leq f_3 \leq \dots \leq f_{n^2}$$

Representación de intensidad de niveles gris

donde $\{f_1, f_2, f_3, \dots, f_{n^2}\}$ son los valores de intensidad o nivel de gris del subconjunto de píxeles en la imagen, que están en la ventana W .

Filtro de la mediana

La mediana M de un conjunto de valores es tal que la mitad de los valores del conjunto son menores que M y la mitad de los valores son mayores que M [1]. Con el objeto de realizar

un filtrado de la mediana en una imagen en un entorno de vecindad, ordenamos las intensidades de la vecindad como se ha explicado, determinamos la mediana y asignamos ésta a la intensidad del píxel central de la ventana W .

Filtro de la moda

De entre todos los valores en el entorno de la vecindad se elige el valor más frecuente, es decir el valor que más veces se aparece. El filtro de la moda es útil para eliminar ruido de tipo impulso y no es tan eficaz para eliminar otros tipos de ruido por ejemplo el Gaussiano.

Filtro de la media

Los filtros de la media trabajan mediante la determinación del promedio del entorno de vecindad $n \times n$ de la ventana. El más básico es el filtro de la media aritmética, que calcula la media aritmética de los píxeles en la ventana, como sigue:

$$Ma = \frac{1}{nm} \sum_{(x,y) \in nm} f(x, y)$$

Formula 1 Formula de la media

donde nm es el número de píxeles en la ventana W de dimensiones $n \times m$. Este filtro suaviza las variaciones dentro de la imagen.

Extracción de regiones

Las operaciones orientadas a la región transforman a la imagen modificando un píxel a la vez y toman en cuenta para dicha transformación los píxeles vecinos. Dicha transformación se puede aplicar a toda la imagen o a una región de ella.

Diremos que los píxeles vecinos de primer orden son aquellos contiguos a él, en una retícula cartesiana regular un píxel, digamos aquel ubicado en la coordenada (x, y) tiene 8 primeros vecinos, si denotamos por $f(x, y)$ al píxel de referencia en la figura 3 se muestran estos.

Figura 3: Primeros vecinos del píxel $f(x, y)$

$I[i-1, j-1]$	$I[i, j-1]$	$I[i+1, j-1]$
$I[i-1, j]$	$I[i, j]$	$I[i+1, j]$
$I[i-1, j+1]$	$I[i, j+1]$	$I[i+1, j+1]$

Muchos filtros regionales utilizan de uno a ocho píxeles vecinos, en particular podemos decir que una transformación regional simple que involucra a los primeros vecinos en una transformación de la forma:

$$I'[i, j] = F(I[i + \alpha, j + \beta]), \alpha, \beta \in \{-1, 0, 1\}$$

Formula 2 Formula de transformación regional

Existen también transformaciones que consideran vecinos más lejanos.

Binarización

La binarización es una de las técnicas más antiguas de procesamiento de imágenes ya que en el pasado, el primer analizador de imágenes automatizado solamente trabajaba con imágenes que estuvieran binarizadas[1,6].

Este método se utiliza debido a tres motivos principales:

- Permite el ahorro de memoria, ya que las imágenes binarizadas son 8 veces más pequeñas que las imágenes en blanco y negro.
- Con las imágenes binarizadas se pueden detectar objetos que se encuentran en la imagen.
- Algunas transformaciones solamente se pueden realizar a cabo con imágenes que se encuentren binarizadas [4].

La binarización selecciona los píxeles que tienen un valor particular que se encuentren dentro de un rango específico. Esto sirve para encontrar objetos que se encuentren dentro de la imagen si se conoce su nivel de brillo. Hay dos principales métodos para realizar la binarización: la *Uniforme* y la *Adaptable*.

Binarización Uniforme

En este método, se buscan los píxeles que se encuentran debajo de un nivel específico, se considerarán como píxeles color blanco y los píxeles que se encuentran por arriba del nivel de grises que contiene la imagen [6].

Binarización Óptima

Este método usualmente busca el valor del umbral que separe la imagen de su fondo, lo que significa que la intensidad de color de la imagen es mayor a la del fondo de la misma lo que proporciona un valor que sea el óptimo para realizar esta separación. Una de las técnicas más utilizadas que implementan la Binarización Óptima es el método de *Otsu's*, el cual maximiza la probabilidad que el umbral que es elegido logra una separación de la imagen entre los objetos y el fondo. Esto es logrado mediante la selección de un umbral que proporcione una mejor separación de clases, para todos los píxeles en la imagen. La base de este método es la utilización del histograma normalizado, donde el número de puntos en cada nivel es dividido entre el número de puntos totales de la imagen. Así, la distribución de probabilidad de los niveles de intensidad está dada por: [4]

$$P(l) = \frac{N(l)}{N}$$

Formula 3 Formula de calculo distribución de probabilidad

Donde:

$P(l)$ =Distribución de probabilidad

$N(l)$ =Número de puntos por nivel de la imagen nueva

N =Puntos de entrada

Esto se puede utilizar para calcular los momentos acumulativos del cero-y-primer orden del histograma normalizado hasta el k-ésimo nivel mediante: [4]

$$w(k) = \sum_{l=1}^k p(l)$$

Formula 4 Formula de momentos acumulativos

Donde:

$P(l)$ =Distribución de probabilidad por nivel

K =Número máximo de niveles

y

$$\mu(k) = \sum_{l=1}^k l * p(l)$$

Formula 5 Formula de clases

Donde:

$P(l)$ =Distribución de probabilidad por nivel

K =Número máximo de niveles

La media del nivel total de la imagen está dada por: []

$$\mu T = \sum_{l=1}^{N \max} l * p(l)$$

Formula 6 Formula de la media del nivel total

Donde:

$N \max$ =Brillo máximo de la imagen nueva

$P(l)$ =Distribución de probabilidad por nivel

La varianza de la separación de clases está dada por la relación: [4]

$$\delta^2_B(k) = \frac{(\mu T * w(k) - \mu(k))^2}{w(k)(1 - w(k))} \forall k \in 1, N \max$$

Formula 7 Formula de varianza de separación de clases

Donde:

μT =Media del nivel total de la imagen

$w(k)$ =Momentos acumulativos

$\mu(k)$ =Momentos acumulativos

K =Nivel

El umbral óptimo es el nivel en el cual la varianza de separación de clases está en su máximo, llamado el umbral óptimo T_{opt} dada por: [4]

$$\delta_B^2(T_{opt}) = \max_{1 \leq k \leq \max} (\delta_B^2(k))$$

Formula 8 Formula de umbral optimo

Donde:

K =Nivel

T_{opt} =Umbral Óptimo

$\delta_B^2(k)$ =Varianza del nivel

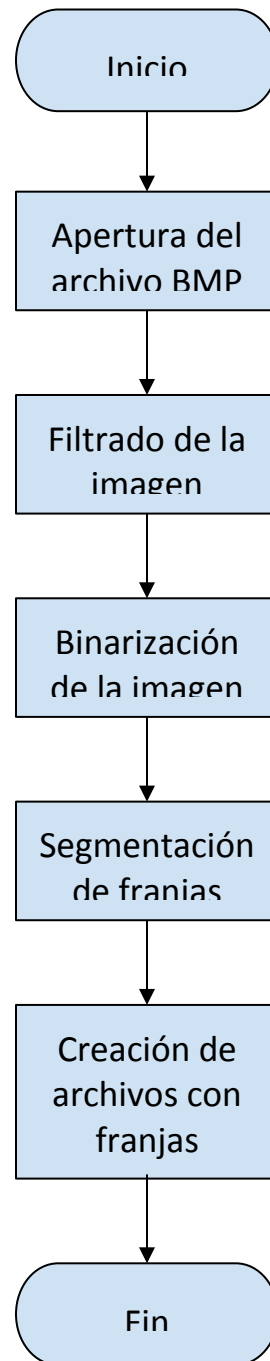
Una comparación entre el método uniforme y el algoritmo óptimo utilizado por el método Otsu's muestra que el método de *Otsu's* considera un umbral más pequeño que se puede utilizar en el método uniforme, sin embargo, el método de *Otsu's* se puede realizar de manera automática, lo cual deja a un lado la realización de pruebas para encontrar el mejor umbral que se ajuste a lo que se está realizando.

También existe el método *local adaptable* la cual es utilizada comúnmente para binarizar documentos con el propósito de localizar caracteres específicos. Esta técnica es utilizada comúnmente para el reconocimiento de patrones estáticos; el objeto a binarizar es clasificado de acuerdo con propiedades estáticas. [4]

Metodología

Con el fin de lograr el objetivo de la detección y posterior segmentación de las franjas de un patrón de luz estructurada se trabajó con imágenes de tipo BMP con una profundidad de 8 bits por pixel. La figura 4 muestra el proceso realizado.

Figura 4: Algoritmo para detección y segmentación de franjas



Para la apertura de un archivo BMP se trabajó con el lenguaje de programación de propósito general C++ el cual es orientado a objetos y una de sus fortalezas es la velocidad

de procesamiento en comparación con algunos otros lenguajes del mismo estilo. Un archivo de este tipo tiene dos secciones: La cabecera y la información de la imagen, la primera de ellas proporciona los datos necesarios para trabajar con el resto del archivo (figura 5).

Figura 5: Estructura de datos para leer la cabecera de un archivo

```

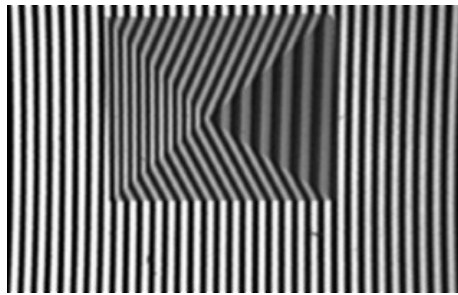
typedef struct {
    unsigned char b;
    unsigned char m;
    unsigned int tam;
    unsigned short reservado1;
    unsigned short reservado2;
    unsigned int inicio_datos;
    unsigned int tam_cabecera;
    int ancho;
    int alto;
    unsigned short numero_planos;
    unsigned short tamano_punto;
    unsigned int compresion;
    unsigned int tamano_imagen;
    unsigned int res_horizontal;
    unsigned int res_vertical;
    unsigned int tabla_colores;
    unsigned int colores_importantes;
} Cabecera;

```

En la figura 5 se muestra la estructura de datos usada para leer la información de la cabecera del archivo que contiene la imagen con el patrón de franjas, en el cual los campos que son útiles son el “inicio_datos”, el cual sirve para desplazarse dentro del archivo a la posición exacta donde inicia la imagen. Los campos de “ancho” y “alto” determinan el tamaño de la matriz en la cual se leerá los diferentes niveles de gris de la imagen. El campo de “tamano_punto” es el que contiene el tamaño de bits por pixel.

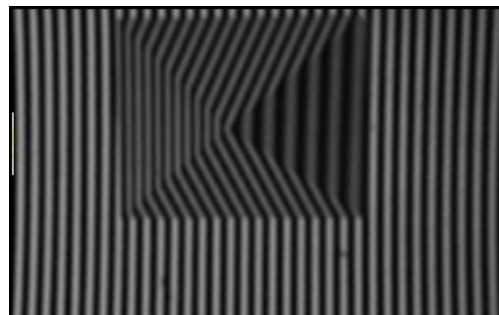
Un ejemplo de archivo con el que se trabajó lo muestra la figura 1.6, el cual tiene unas dimensiones de 200 x 320 con una profundidad de 8 bits.

Figura 6: Imagen con franjas de profundidad de 8 bits



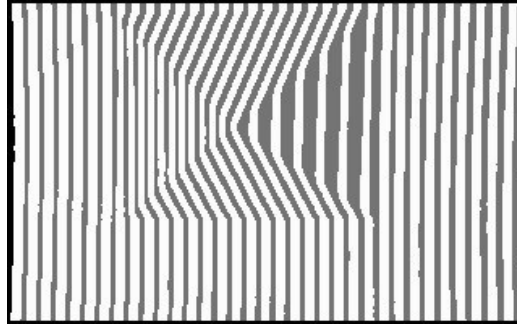
Con la información almacenada en una matriz que contiene los diferentes niveles de intensidad de los píxeles de la que está compuesta la imagen. Se procedió a aplicar un filtro de Media de tamaño de ventana 5 X 5 con un umbral de 65, esto permite eliminar el ruido que existe en los valores de la matriz, este ruido se puede producir por diversos factores entre ellos la forma de cómo se capturan las imágenes, factores ambientales, etc. El resultado de la aplicación del filtro se muestra en la figura 7.

Figura 7: Imagen con aplicación de filtro de la media



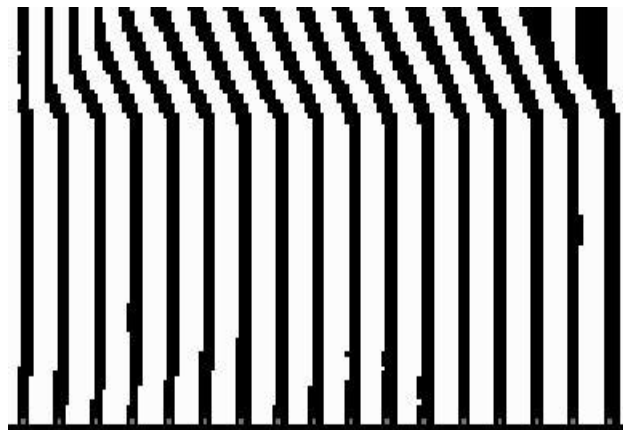
Con los valores filtrados en la matriz se procede a binarizar la imagen lo que facilitará la segmentación de las diferentes franjas de que se compone el patrón, como lo muestra la figura 8.

Figura 8: Imagen binarizada



Con valores en la matriz de 0 y 255 en todos los valores x , y que pertenecen al ancho y alto de la imagen se procede a buscar las bases de las diferentes franjas en la imagen, lo cual permite hacer el seguimiento de la franja de luz dentro de la imagen (Figura 9). Para lograr el seguimiento se implementó un algoritmo de código de cadena que permite dar seguimiento a los píxeles que componen la franja de luz, ya con esto se procede a formar archivos tipo BMP con solo una franja de luz en él.

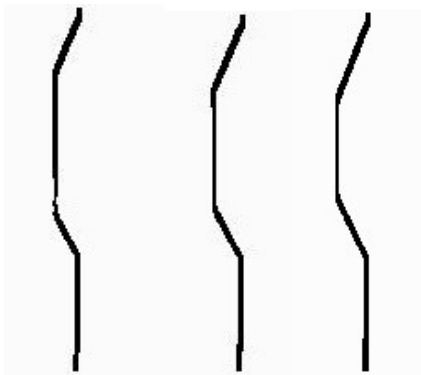
Figura 9: Imagen con franja de luz



Conclusiones

Con los procesos descritos anteriormente se puede detectar y segmentar de forma efectiva las franjas de luz, en el caso de la imagen de prueba se generan 34 archivos BMP, uno por cada línea que existe en la imagen original. Lo anterior facilitará los procesos siguientes con la finalidad de lograr la recuperación la forma del objeto bajo la luz estructurada (Figura 10).

Figura 10: Ejemplos de imágenes segmentadas



Con el resultado de este trabajo se cuenta ya con una línea por archivo BMP, esto con la finalidad de poder hacer un proceso de dilatado o erosión, lo que permitirá contar con una línea de un pixel de ancho, esto con el objetivo de poder muestrear con una determinada frecuencia para lograr contar con los centros de sirvan de base para algún método de interpolación que se seleccione para poder reproducir la forma del objeto.

Referencias:

1. G. Pajares and J M de la Cruz García, *Visión por computador*, Alfaomega Ra-Ma. 1-3. 2002.
2. A. Rosenfeld and A. C. Kak, **Digital Picture Processing**. Academic Press, 1982.
3. R. C. González y R. E. Woods. **Tratamiento digital de imágenes**. Addison-Wesley/Díaz de Santos. 6-7.1996.
4. M. Martín Ortiz, **Procesamiento Digital de Imágenes**. FCC-BUAP-Rev., 2013
5. González Marcos A., **Técnicas y algoritmos básicos de Visión artificial**, Universidad de la Rioja, 2006
6. D. Torres Armenta, **Digitalización de objetos a través de técnicas de proyección de luz estructurada y reconstrucción mediante técnicas de computación suave**, CIO REV-2010