

RADON, HOUGH, ACUMULACIÓN Y EL MÉTODO SDR

k. Voss, H. Suesse and W. Ortmann

Comunicación Técnica No I-04-05/05-07-2004
(CC/CIMAT)



RADON, HOUGH, ACUMULACIÓN Y EL MÉTODO SDR

K. Voss, H. Suesse, W.Ortmann
Universidad Friedrich Schiller, Jena / Alemania
Departamento de Matemáticas y Computación
nbs@uni-jena.de , www.pandora.inf.uni-jena.de

Trabajo realizado en parte durante de la estancia
del profesor K.Voss en el CIMAT (de enero hasta junio 2004)

Resumen

En este trabajo se investiga el método de acumulación como generalización de la transformadas de Radon y de Hough y como mejoramiento del método RANSAC.

Mostramos que tanto el ajuste de conjuntos de puntos para lograr figuras geométricas (líneas rectas, círculos, elipses) como la determinación de transformaciones (traslaciones, movimientos, transformaciones afines) se pueden realizar a partir de este nuevo método de acumulación.

En caso que no se pueda determinar puntos correspondientes con una precisión suficiente usamos el método SDR (*shift detection by restoration*). Este método considera un desplazamiento entre dos imágenes I_1 y I_2 como una convolución $I_2 = D * I_1$ con D como la imagen de un impulso unitario desplazado.

La inversión formal $D = I_2 * I_1^{-1}$ de esta ecuación se reemplaza por una solución obtenida a base de una pseudo-inversa (método de restauración). Esta solución proporciona buenos resultados también para imágenes muy perturbadas y para grandes desplazamientos.

En caso de transformaciones afines entre dos imágenes se pueden determinar desplazamientos locales en pequeñas ventanas y a partir de esta lista de los vectores de translación se calculan los coeficientes de la transformación afín.

Índice General

1	Introducción	3
2	Métodos de acumulación	4
2.1	Transformada de Radon	4
2.2	Transformada de Hough	5
2.3	Método RANSAC	7
2.4	Método de la acumulación	8
3	Procedimientos secundarios	9
3.1	Método de la proyección	9
3.2	Método de la agrupación	11
4	Detección de figuras más complejas	11
4.1	Acumulación para círculos	11
4.2	Acumulación para elipses	13
4.3	Acumulación para cuadrados	14
5	Transformaciones geométricas	17
5.1	Traslaciones	17
5.2	Transformaciones homotéticas	19
5.3	Transformaciones afines	20
6	Detección de desplazamientos	22
6.1	Convoluciones y operadores lineales	22
6.2	Inversión de convoluciones	24
6.3	El método SDR	26
6.4	Investigaciones experimentales	29
7	Conclusiones	32
8	Referencias	32

1 Introducción

El reconocimiento de líneas rectas es una tarea fundamental en el campo de procesamiento de imágenes. Por solucionar este problema, se usa la transformada de Hough desde hace cuarenta años (Hough, 1962): para cada punto $q=q(x,y)$ de un conjunto Q de puntos se acumulan las parejas (p,φ) de los valores discretos de la curva $p=x\cos\varphi+y\sin\varphi$ en el espacio de parámetros (p,φ) y entonces se detectan los picos más destacados en este espacio (ver, por ejemplo, Gonzalez y Woods, 1993; Pratt, 1991; Sonka *et al.*, 1993).

Se puede generalizar la idea fundamental de la transformada de Hough también para figuras geométricas con más de dos parámetros (Ballard, 1981). Pero en aquel enfoque se calculan primeramente unas relaciones geométricas para el modelo, se almacenan los valores de estas relaciones en tablas de consulta y se comparan finalmente las relaciones geométricas aparecidas en las imágenes con los valores contenidos en las tablas.

Contrariamente a este método de la extensión de la transformada de Hough, la generalización directa usa los parámetros geométricos de una figura. Por ejemplo, se describe un círculo por $(x-c)^2+(y-d)^2=r^2$. Con valores x y y fijados, esta ecuación determina un cono doble en el espacio tridimensional de las coordenadas (c,d,r) . Si acumulamos estos conos en el (c,d,r) -espacio para todos de los puntos (x,y) , entonces los picos más altos en el espacio tridimensional caracterizan los círculos en el (x,y) -espacio bidimensional.

No conocemos ninguna publicación que investiga la acumulación de conos dobles en un espacio tri-dimensional. Por eso queremos describir un método general de acumulación que trabaja en espacios con más dimensiones (3 para círculos, 5 para elipses, 4 para cuadrados).

Este método es aplicable también en el caso de la determinación de transformaciones geométricas. Usando dos listas de puntos (conteniendo puntos correspondientes y puntos de perturbación) tenemos que determinar los parámetros de una traslación (2 parámetros), de un movimiento (3 parámetros) o de una transformación afín (6 parámetros).

Algunos enfoques anteriores usan también la idea de los parámetros de figuras o transformaciones. Pero no se explotan consecuentemente los beneficios de la acumulación (método RANSAC, ver Fischler y Bolles, 1981) o se emplean métodos de aproximación (como el uso de los puntos de la cubierta convexa, ver Goshtasby y Stockman, 1985).

Este trabajo está organizado de la siguiente manera: Como primer paso investigamos los métodos de Radon y de Hough, el método RANSAC y el enfoque de la acumulación por medio de la detección de líneas rectas. Entonces - para mejorar los métodos para la práctica - describimos unos procedimientos como el método de la proyección y de la agrupación directa. En el parte siguiente detectamos círculos, elipses, cuadrados y por último tratamos la detección de las transformaciones geométricas.

Finalmente, en la sexta parte introducimos el método SDR (*shift detection by restoration*) que se basa en la inversión de convoluciones. A partir de este método se pueden determinar listas de puntos correspondientes en dos imágenes sin conocimientos sobre el tipo o el contenido de las imágenes. Usando estas listas es posible calcular transformaciones especiales (movimientos, transformaciones afines) entre ambas imágenes.

2 Métodos de acumulación

2.1 Transformada de Radon

En la practica algunas veces aparecen problemas como los mostrados en la figura 1, es decir queremos detectar líneas rectas en un fondo ruidoso.

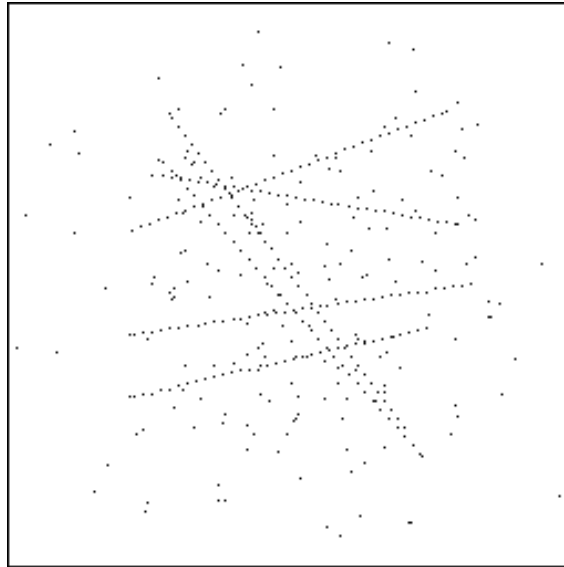


Figura 1: Ejemplo de una imagen con 6 líneas rectas

La solución de tareas de este tipo se puede realizar a partir de la transformación de Radon o de la transformación de Hough. La transformación de Radon fue introducida en matemáticas en 1917 (Radon, 1917, ver también Anger y Portenier, 1992; Deans, 1983). Se integran o suman valores de una función bidimensional $f(x,y)$ a lo largo de una línea recta que es desrita por la forma normal:

$$R_f(p, \varphi) = \int_{x \cos \varphi + y \sin \varphi = p} f(x, y) dx dy$$

Es claro que en la imagen de la figura 1 para las líneas "correctas" esta integral tiene un valor mayor que para otras líneas, donde las integrales proporcionan valores menores. Si interpretamos estos valores como valores de gris de una imagen (el eje de la abscisa describe los ángulos $0 < \varphi < 180^\circ$, y el eje de las ordenadas describe los distancias $-s < p < +s$ con s como el tamaño de la imagen original), entonces obtenemos las imágenes de la figura 2 (la distancia p se refiere al centro de las imágenes).

Las líneas correctas impresionan por los picos marcados en las funciones de valores de gris de las imágenes de Radon. Si hemos determinado más o menos precisamente los lugares (p, φ) de estos picos, entonces podemos seleccionar todos los puntos (x, y) de las imágenes de la figura 1 con una distancia pequeña

$$d = |x \cos \varphi + y \sin \varphi - p|$$

a la línea provisional (quiza con pesos inversamente proporcionales a las distancias). Al final obtenemos, por el método de mínimos cuadrados (o por la optimización lineal, ver Voss y Suesse, 2001) las líneas rectas precisas.

Para calcular la imagen de Radon I_R se cuentan los puntos a lo largo de la línea $L_{p,\varphi}$ contenidos en una franja del ancho δ . Con N como el número de todos los puntos y D como el número medio de las células discretas de las coordenadas p y φ , se obtiene una complejidad $O(ND^2)$ para la determinación de la imagen de Radon (la búsqueda de los picos más altos se puede lograr durante la construcción de I_R)

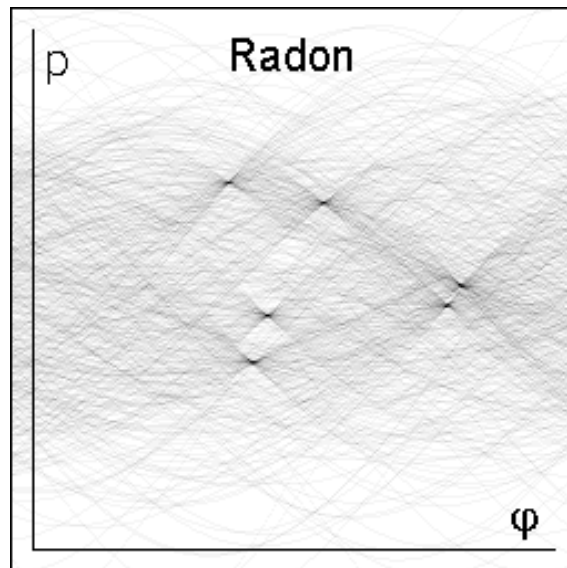


Figura 2: Imagen de Radon de las 6 líneas

2.2 Transformada de Hough

Por desgracia no tenemos transformadas del tipo Radon para otras figuras como círculos o elipses. Pero en 1962 Hough describió un enfoque por el que se puede extender la transformada de Radon a otras figuras y se puede reducir la complejidad (Hough, 1962).

Hough interpretó la ecuación $p = x \cos \varphi + y \sin \varphi$ como una ecuación con parámetros (x, y) y variables (p, φ) en lugar de una ecuación con parámetros (p, φ) y variables (x, y) . Todas las parejas (p, φ) que cumplen la ecuación $p = x \cos \varphi + y \sin \varphi$ para coordenadas dadas x y y describen líneas rectas que pasan por el punto (x, y) .

Estas parejas (p, φ) describen en el espacio bidimensional de las variables p y φ una curva $p = x \cos \varphi + y \sin \varphi$ (una curva de tipo seno o coseno). Ahora se acumulan todas estas curvas en el espacio de las variables p y φ (el "espacio de acumulación" o "espacio de Hough"). El resultado es mostrado en figura 3.

La imagen de Hough tiene aproximadamente la misma estructura que la imagen de Radon, es decir $I_H \sim I_R$. El razon de esta igualdad es la siguiente: Usando el método de Radon contamos todos los puntos (x,y) contenidos en la región $R_{p,\varphi,\delta}$, es decir $(x,y) \in R_{p,\varphi,\delta}$. A partir del método de Hough acumulamos todas las líneas (p,φ) que pasan por el punto (x,y) en una célula y que satisfacen aproximadamente la ecuación de la línea, es decir otra vez $(x,y) \in R_{p,\varphi,\delta}$.

Pero la acumulación en el espacio de Hough necesita sólo una complejidad $O(ND)$ porque se puede determinar para cada uno de los N puntos (x,y) y cada uno de los D ángulos φ el parámetro p , así que se puede aumentar el contenido de la célula (p,φ) .

También se pueden aplicar el método de Hough a otros problemas como la detección de círculos (ver también sección 4.1): A través de un punto (x,y) pasan todos los círculos con centro (c,d) y radio r que cumplen la ecuación $(x-c)^2 + (y-d)^2 = r^2$.

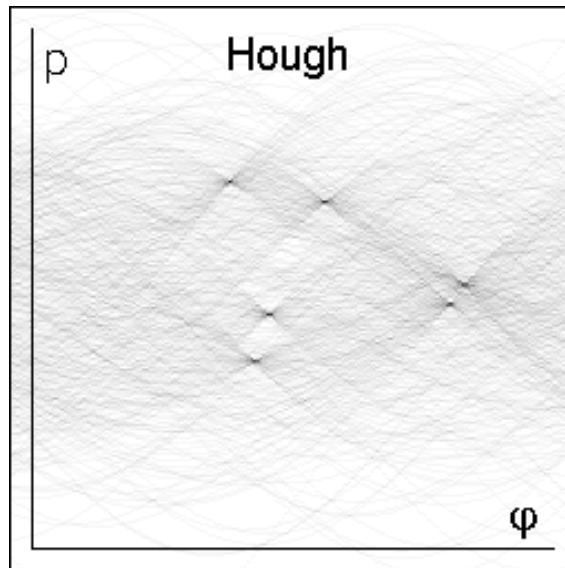


Figura 3: Imagen de Hough de las 6 líneas

Esta ecuación describe un cono doble en el espacio tridimensional de las coordenadas c,d,r . Se tienen que acumular todos los conos determinados por los puntos (x,y) en el espacio (c,d,r) y entonces se tienen que detectar los picos en este espacio. La complejidad de este enfoque es $O(ND^2)$.

En caso de que el radio es conocido, se tiene que solucionar el problema restringido, es decir se tienen que acumular círculos con radio r en un (c,d) -espacio bidimensional (la complejidad es $O(ND)$).

Por otro lado, se puede usar la transformada generalizada de Hough (Ballard, 1981) que trabaja a partir de rasgos típicos de las figuras de interés. Por ejemplo, se usa el método de gradiente para determinar primeramente el centro del círculo como punto de intersección común de las líneas rectas de gradiente, y entonces se determina el radio (Guil y Zapata, 1997; Kimme *et al.*, 1975). Otros autores usan también enfoques simplificados del método original de Hough para la búsqueda de círculos, sea por aceleración de la computación o por superación de las dificultades de la dimensión tridimensional (Davies, 1990; Cao y Deravi, 1990; Li *et al.*, 1986).

No conocemos trabajos que usen el enfoque directo de la generalización del método de Hough, es decir la acumulación de conos dobles en el (c,d,r) -espacio tridimensional.

2.3 Método RANSAC

Una idea nueva para la determinación de figuras y transformaciones es el método RANSAC (*random sample consensus*, Fischler y Bolles 1981). Los autores Fischler y Bolles han formulado su idea como siguiente:

“Given a model that requires a minimum of n data points to instantiate its free parameters, and a set of data points P such that the number of points in P is greater than n , randomly select a subset S of n data points from P and instantiate the model M . Use the instantiated model to determine the subset S^* of points in P that are within some error tolerance. The set S^* is called the consensus set of S .

If the number n_s of points in S^* is greater than some threshold t , use S^* to compute (possibly using least squares) a new mode M^* . If n_s is less than t , randomly select a new subset S^* and repeat the above process. If - after some pre-determined number of trials - no consensus set with t or more members has been found, either solve the model with the largest consensus set found, or terminate in failure.”

Este método busca líneas que tienen muchos puntos en sus vecindades. Pero a diferencia del método de Radon, el método RANSAC busca las líneas aleatoriamente.

Para eso se eligen dos puntos al azar, se determinan la línea recta que pasa estos dos puntos y se cuentan todos los puntos (x,y) que están situados en el región $R_{p,\varphi,\delta}$, es decir $(x,y) \in R_{p,\varphi,\delta}$. Si el número de los puntos es suficiente (mayor de un umbral o un límite), entonces se decide que una línea correcta se ha hallado.

Claro, el método de RANSAC busca aleatoriamente picos en la imagen de Radon (ver figura 2). La efectividad de esta búsqueda depende de la probabilidad de que los dos puntos elegidos residan en la línea correcta.

A partir de los dos puntos $p_1=(x_1,y_1)$ y $p_2=(x_2,y_2)$ determinamos la línea recta $L_{p,\varphi}$ por las ecuaciones

$$p = \frac{x_1 y_2 - x_2 y_1}{\sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}}, \quad \varphi = \arctan \frac{x_1 - x_2}{y_1 - y_2}$$

En la figura 4 las probabilidades de nuestro ejemplo de la figura 1 son representadas para cada una de las parejas (p_1,p_2) de los puntos respectivamente para cada uno de los valores (p,φ) correspondientes (por supuesto, el método de RANSAC no calcula todos estos valores).

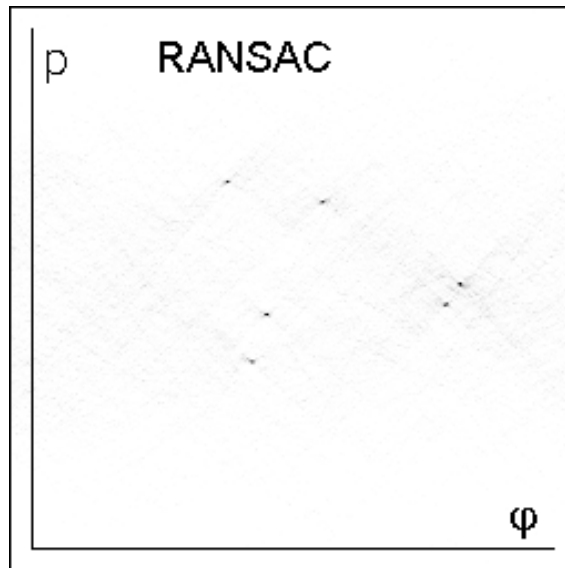


Figura 4: Imagen de RANSAC de 6 líneas

Se puede ver que los picos de la probabilidad en la figura 4 son más destacados que los picos de las imágenes de Radon I_R o de Hough I_H . La razón para este hecho es que en lugar de un punto usamos dos puntos: En las imágenes I_R o I_H determinamos para cada línea $L_{p,\varphi}$ el número de los puntos. Por otro lado, en la imagen I_p de las probabilidades de RANSAC tenemos que contar todas las parejas de puntos que se hallan en una línea $L_{p,\varphi}$, así que logramos $I_p \sim I_R^2 \sim I_H^2$.

En realidad, los cuadrados de los valores de gris de las imágenes I_R o I_H resultan en una imagen como figura 4.

La complejidad de la computación de la imagen de la probabilidad I_p es $O(N^2)$ ya que tenemos $N(N-1)/2$ parejas de puntos. El método de RANSAC termina el cálculo si se ha obtenido una línea con un número suficiente de puntos.

El método original RANSAC no calcula la imagen completa como en la figura 4. En lugar de eso se decide si el número de los puntos de una línea es suficiente por la validez de la línea.

Sin el conocimiento sobre el conjunto de los puntos, no se puede especificar un límite para esta decisión así que la terminación de la búsqueda por líneas correctas puede proporcionar resultados falsos. Por otro lado, la imagen de RANSAC muestra los picos apuntando las líneas correctas más acertadamente que las imágenes de Radon o de Hough.

2.4 Método de la acumulación

Nuestro nuevo método de la acumulación combina las ventajas del método de Radon y del método RANSAC: elegimos aleatoriamente parejas de puntos, determinamos la línea correspondiente, calculamos los puntos vecinos a la línea y acumulamos estos números.

De esta manera resulta la imagen de la acumulación (ver figura 5). Porque acumulamos números distintos, los picos son más marcados otra vez: cada línea aparece con una probabilidad dada por la imagen RANSAC y el incremento durante la acumulación es determinado por la imagen de Radon. Por eso obtenemos el resultado $I_A \sim I_P \sim I_H^2 \sim I_R^3$.

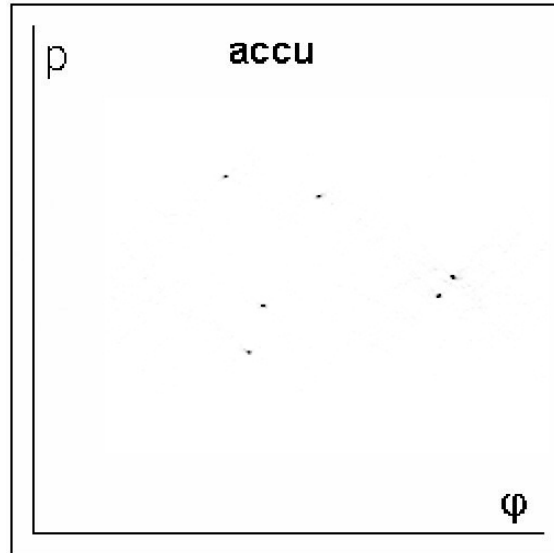


Figura 5: Imagen de la acumulación total de las 6 líneas

La complejidad por determinar la imagen I_A es $O(N^3)$. Por eso tenemos que buscar métodos mejorados y más efectivos. La idea adecuada ya está dada por el método RANSAC: se busca aleatoriamente por líneas (es decir, por parejas de puntos) y se acumulan los resultados. De esta manera no es necesario investigar todas las $N(N-1)/2$ parejas, sino que los picos ya aparecen si se han investigado N o $const \cdot N$ parejas.

En caso de las tareas más complicadas (determinación de círculos, de elipses, de transformaciones afines) la acumulación necesita un espacio de 3 o 5 o 6 dimensiones, es decir una memoria de 10^6 - 10^{12} byte. Se puede superar esta problema por el método de la proyección (ver sección 3.1) o por el método de la “agrupación en línea” (ver sección 3.2).

3 Procedimientos Secundarios

3.1 Método de la proyección

En este parte se investiga la acumulación por el método de proyección. Este método es necesario ya que se no puede mostrar el resultado de la acumulación para espacios con más de dos dimensiones. Pero a partir la nitidez de los picos en la imagen I_A es posible proyectar el espacio de dimensión alta a un espacio de menos dimensiones.

Por ejemplo, podemos proyectar las imágenes I de las figuras 2-5 al eje p (la ordenada) o al eje φ (la abcisa). De esta manera se produce un histograma H de los resultados de la acumulación (ver figura 6). Podemos escribir esta operación por el operador \mathbf{P} de la proyección como $H=\mathbf{P}I$.

Claro, en el histograma de Hough contamos todas las curvas acumuladas para cada valor del ángulo φ y por eso obtenemos casi una función constante (y obtenemos también $H_R \sim H_H$). En el histograma RANSAC $H_P = \mathbf{P}I_P$ los picos son más o menos destacados, y al final no hay dificultades en la detección de los seis picos del histograma $H_A = \mathbf{P}I_A$ de la acumulación.

Por supuesto, podemos calcular las imágenes $I_R = g_R(p, \varphi)$ o $I_H = g_H(p, \varphi)$ y entonces determinar los histogramas H_R y H_H , por ejemplo

$$H_R(\varphi) = \sum_{\varphi} (g_R(p, \varphi))^2$$

Pero en este caso podemos determinar primeramente la imagen de Radon con una complejidad de $O(ND^2)$, es decir, la construcción del histograma H_R necesita también la complejidad temporal $O(ND^2)$ en caso de la imagen I_R o la complejidad $O(ND)$ en caso de la imagen I_H así como $O(D^2)$ células de memoria.

Al contrario, la determinación directa de la imagen I_P por el método de RANSAC necesita solamente una complejidad temporal $O(N^2)$ y sólo $O(D)$ células de memoria. En caso del histograma H_A de la acumulación necesitamos $O(N^3)$ operaciones.

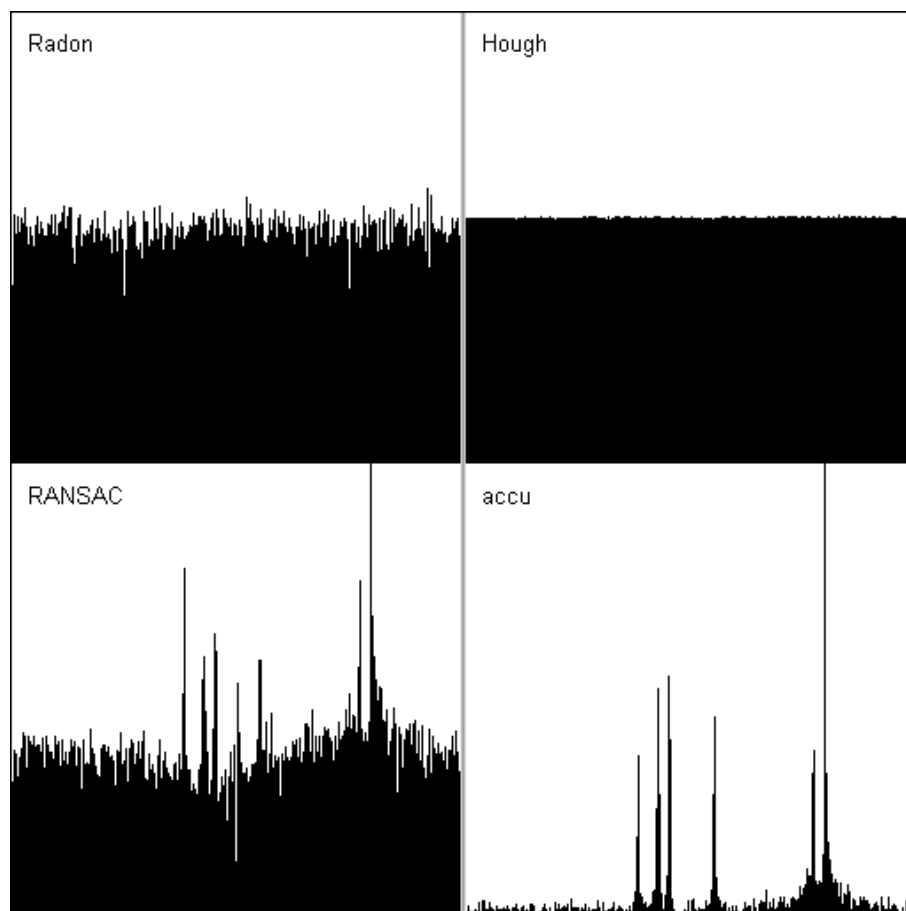


Figura 6: Imágenes de la proyección al eje del ángulo φ

3.2 Método de la agrupación

El método de RANSAC y el método de la acumulación proporcionan buenos resultados para las imágenes y para los histogramas. Pero tanto la determinación de las imágenes como la determinación de los histogramas requieren mucho gasto de tiempo y de memoria. Sin embargo hay una posibilidad para superar estas dificultades: por el enfoque de la agrupación (*clustering*) en línea podemos ahorrar tiempo y memoria.

El valor v de una medida de calidad depende del problema actual. En caso de un círculo no podemos usar el número de los puntos que están situados cerca del círculo ya que entonces círculos pequeños tienen valores pequeños de v . Por eso usamos la densidad de los puntos como la medida de calidad, es decir, el número de los puntos vecinos del círculo dividido por el perímetro del círculo.

En caso de una elipse (ver sección 4.2) es muy difícil determinar la distancia de un punto a la elipse (ver Voss y Suesse 1995, Voss *et al.* 1999, Voss *et al.* 2002). Es más simple a contar todos los puntos que satisfacen la ecuación del segundo orden de la elipse hasta un pequeño error y a referir este número al perímetro de la elipse (aproximadamente $\pi(a+b)/4$ con a y b como las longitudes de los ejes de la elipse).

4 Detección de figuras más complejas

4.1 Acumulación para círculos

En caso de la detección de un círculo hay cuatro posibilidades para solucionar el problema. Un círculo es descrito por la ecuación

$$(x - c)^2 + (y - d)^2 = r^2$$

y por eso la “imagen de Radon” es una imagen tridimensional o un espacio tridimensional de las variables c, d y r . El “valor de gris” en el espacio de Radon es la densidad de los puntos vecinos del círculo. La determinación de la imagen tridimensional de Radon tiene una complejidad de $O(ND^3)$.

La segunda posibilidad es el cálculo de la imagen tridimensional de Hough. En este caso investigamos todos los círculos (c, d, r) que pasan un punto fijo (x, y) . La ecuación describe un cono doble en el (c, d, r) -espacio que es fijado por los parámetros x y y . Se tienen que acumular todos estos N conos dobles en el espacio tridimensional y que buscar por el pico máximo de la acumulación.

El enfoque tercero (no usado en la práctica del procesamiento de imágenes) usa dos puntos (x_1, y_1) y (x_2, y_2) y investiga todos los círculos que pasan ambos puntos. El sistema de ecuaciones

$$\begin{aligned}(x_1 - c)^2 + (y_1 - d)^2 &= r^2 \\ (x_2 - c)^2 + (y_2 - d)^2 &= r^2\end{aligned}$$

es igual a la ecuación

$$2c(x_2 - x_1) + 2d(y_2 - y_1) = x_2^2 + y_2^2 - x_1^2 - y_1^2$$

Esta ecuación describe una línea recta en el (c,d) -espacio así que tenemos que acumular todas de estas líneas. El pico más destacado señala las coordenadas c,d del centro del círculo. La complejidad de este enfoque es $O(N^2D)$ porque tenemos una línea recta para cada pareja de los N puntos.

Por supuesto, podemos usar los tres puntos por la determinación unívoca de un círculo. Esta cuarta posibilidad está usada por el método de RANSAC pero también en (Cao y Deravi 1990) que emplean la idea de la proyección al (c,d) -plano. Ahora tenemos la complejidad $O(N^3)$.

En la figura 7 vemos tres ejemplos y los resultados del método RANSAC (o también el método de la acumulación). El tiempo por la determinación de los círculos correctos es 20 milisegundos en el primero ejemplo y 150 milisegundos en el tercer ejemplo (procesador de 1 GHz). El número de los puntos circulares es 60, el número de los puntos de perturbación es 300, 600 y 1200 para los tres ejemplos.

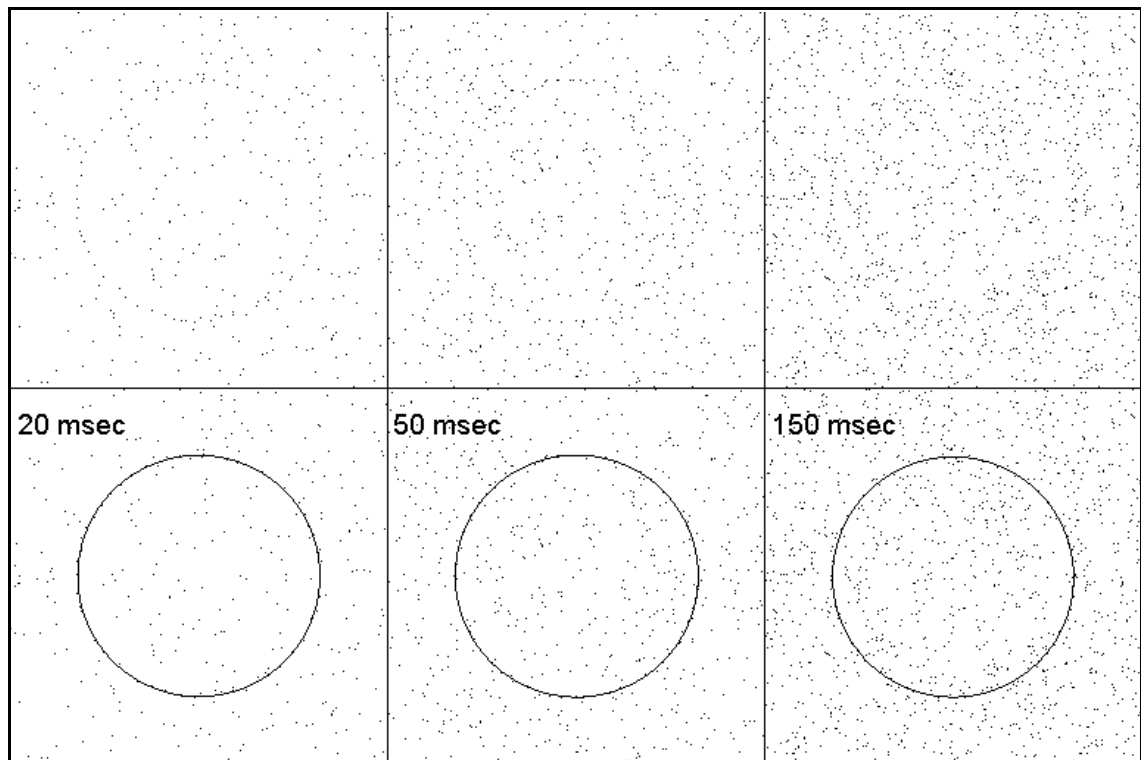


Figura 7: Detección de círculos por el método de acumulación

El valor de la razón de la altura del pico más destacado a la altura del segundo pico. Si este valor es mayor que 2 o 3, entonces tenemos una probabilidad muy alta de la detección del círculo correcto y sólo una parte pequeña de las $O(N^3)$ ternas de puntos se tiene que acumular.

4.2 Acumulación para elipses

Una curva de segundo orden (elipse o hipérbola) del plano euclideo está determinada por la ecuación

$$Ax^2 + Bxy + Cy^2 + Dx + Ey + F = 0$$

Hay algunas posibilidades para reducir el conjunto de los parámetros (ver Rosin 1993). Aquí usamos la condición $A+C=1$ que implica la invariancia del ajuste (con respecto a un conjunto de puntos) contra traslaciones, rotaciones y escalaciones. Por eso, se determinan los parámetros a partir de cinco ecuaciones

$$(1 - C)x_i^2 + Bx_i y_i + Cy_i^2 + Dx_i + Ey_i + F = 0$$

para $i=1\dots 5$. Si el discriminante $\delta = (1-C)C - B^2/4$ es positivo, entonces hemos hallado una elipse ($\delta < 0$ caracteriza una hipérbola).

El proceso de la acumulación consta de la determinación de los parámetros $A\dots F$ a partir de (todos o suficientemente muchos) quintuplas de puntos (x_i, y_i) y de la detección de un pico de la acumulación (en el espacio 5-dimensional o por la agrupación en línea).

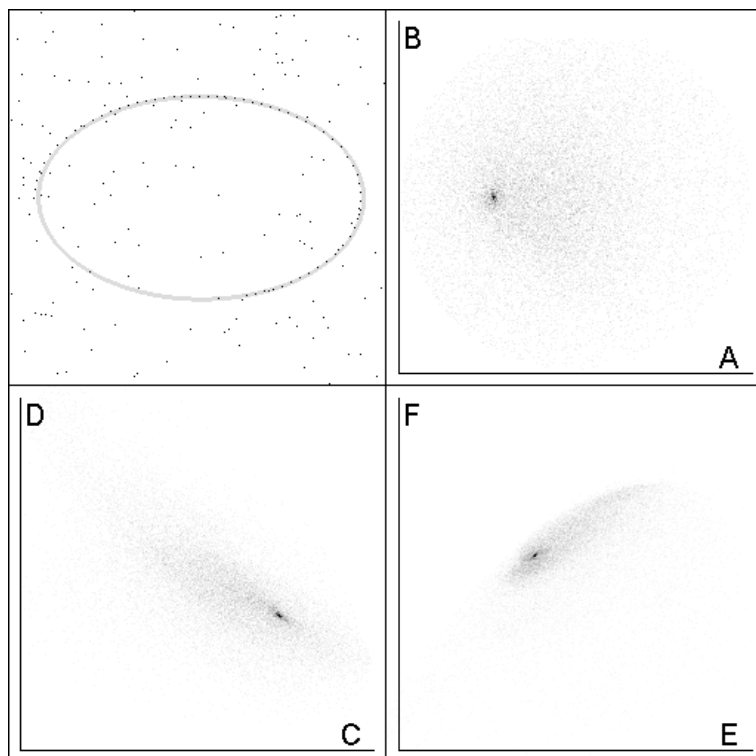


Figura 8: Acumulación y proyección de seis parámetros en planos (A,B) , (D,C) y (E,F)

Una elipse está también definida geoméricamente por su centro (x_0, y_0) , sus ejes (a, b) y el ángulo φ del eje mayor contra la abcisa. Se pueden interpretar estos parámetros geoméricos más sencillamente que los parámetros algebraicos $A\dots F$.

El método de RANSAC con la proyección del espacio de los parámetros a los planos (A,B) , (D,C) y (E,F) proporciona resultados como el mostrado en la figura 8. Los picos de valores grises en las imágenes de la proyección dan los parámetros óptimos A,B,C,D,E,F . La elipse correspondiente está mostrada en el parte arriba a la izquierda de la figura 8.

En caso de la agrupación de los parámetros geométricos x_0, y_0, a, b, φ y la proyección en los planos (a,b) , (x_0, y_0) y (a, φ) se proporciona la figura 9.

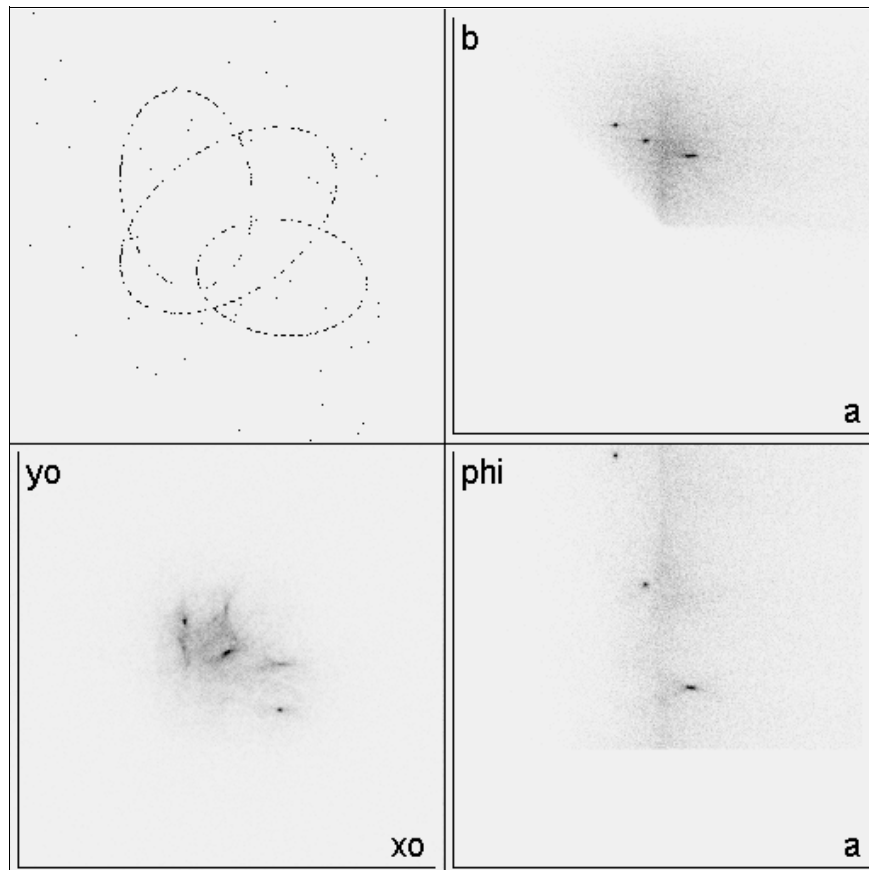


Figura 9: Acumulación de los parámetros x_0, y_0, a, b, φ en tres planos de proyección

4.3 Acumulación para cuadrados

Se puede usar el método de acumulación también si no hay sólo una ecuación. Como ejemplo queremos investigar la detección de los cuadrados. Un cuadrado es determinado por 4 parámetros:

- x_s, y_s las coordenadas del centro
- a la longitud de los lados
- φ el ángulo de los lados (entre 0° y 90°)

Los lados del cuadrado están situados en líneas rectas que son descritas por sus ecuaciones en la forma normal:

$$\begin{aligned} (x_1 - x_s) \cos\varphi + (y_1 - y_s) \sin\varphi &= a/2 \\ (x_2 - x_s) \cos(\varphi + \pi/2) + (y_2 - y_s) \sin(\varphi + \pi/2) &= a/2 \\ (x_3 - x_s) \cos(\varphi + \pi) + (y_3 - y_s) \sin(\varphi + \pi) &= a/2 \\ (x_4 - x_s) \cos(\varphi + 3\pi/2) + (y_4 - y_s) \sin(\varphi + 3\pi/2) &= a/2 \end{aligned}$$

Por la adición de las ecuaciones I+III y de las ecuaciones II+IV se obtiene

$$\begin{aligned} (x_1 - x_3) \cos\varphi + (y_1 - y_3) \sin\varphi &= a/2 \\ -(x_2 - x_4) \sin\varphi + (y_2 - y_4) \cos\varphi &= a/2 \end{aligned}$$

y por eso

$$\tan\varphi = \frac{x_1 - y_2 - x_3 + y_4}{-y_1 - x_2 + y_3 + x_4}.$$

Si el ángulo φ es conocido entonces podemos determinar la longitud a y las coordenadas x_s y y_s . Una condición de la solución es que los puntos 1,2,3,4 son las esquinas de un cuadrilátero convexo.

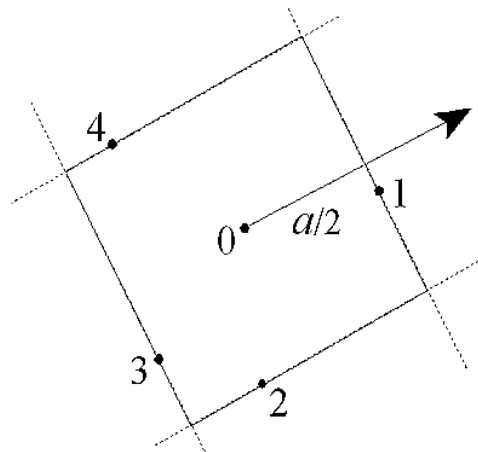


Figura 10: Descripción de un cuadrado por 4 puntos

La condición de que los 4 puntos están situados en los 4 lados de un cuadrado es que los puntos 1 y 3 están entre las líneas rectas L_2 y L_4 y que los puntos 2 y 4 están entre las líneas rectas L_1 y L_3 (ver figura 11).

En la figura 12 se muestra un ejemplo para la detección de un cuadrado. Hay muchos puntos de ruido y también una línea de perturbación. Sin embargo la acumulación de los parámetros en dos planos (x_s, y_s) y (φ, a) nos da picos acertados y por eso la solución correcta.

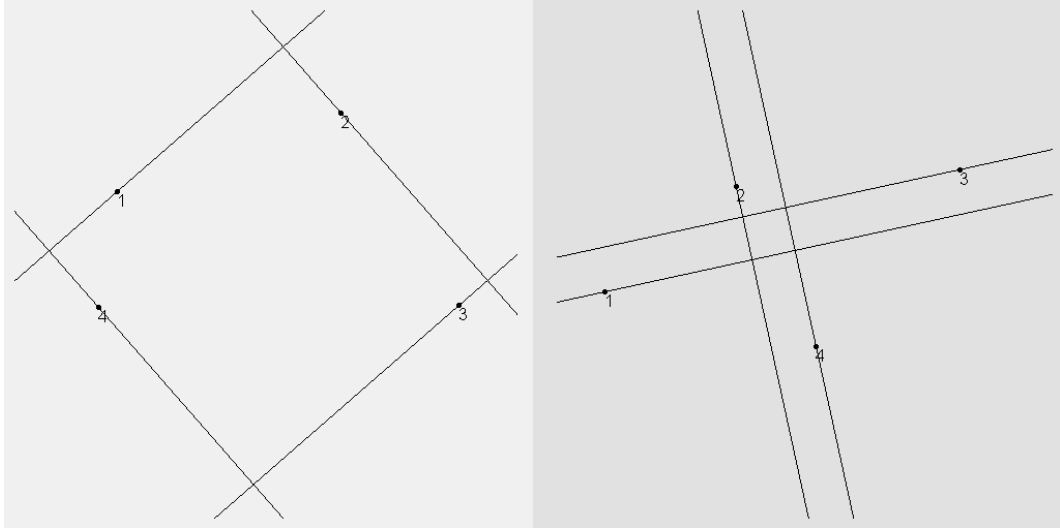


Figura 11: Cuatro puntos en los lados de un cuadrado (izquierda) y en dos parejas de líneas rectas que son mutuamente ortogonales (derecha)

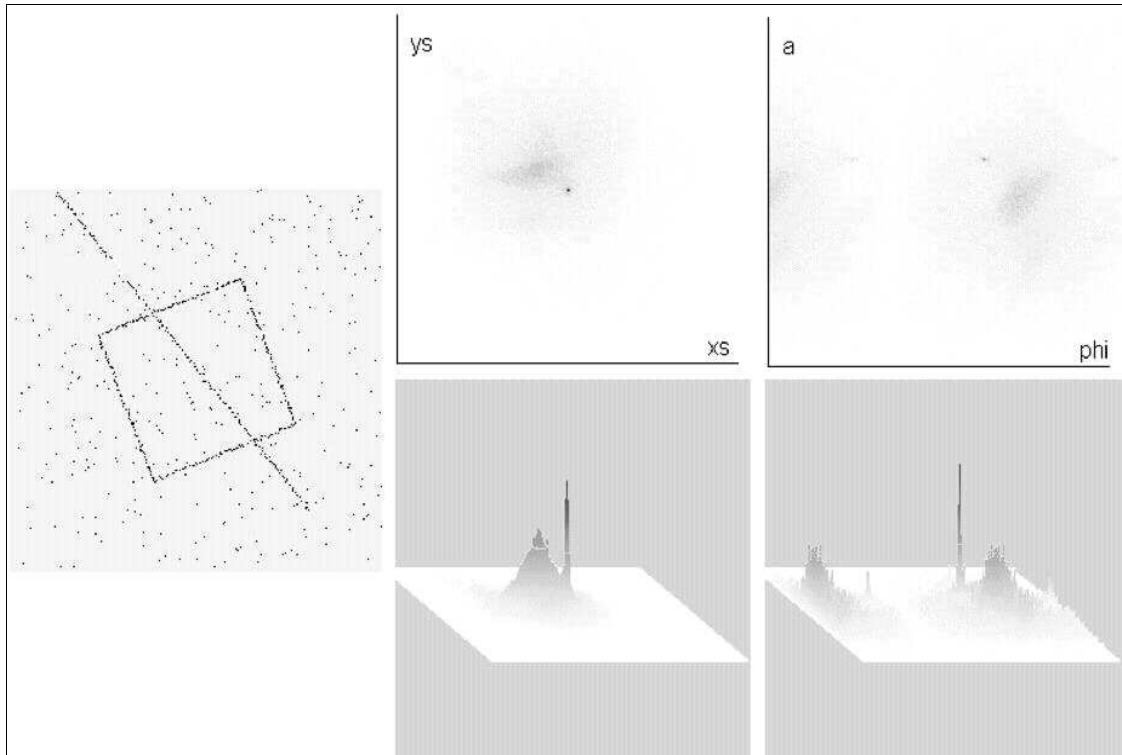


Figura 12: Ejemplo para la detección de un cuadrado por el método de acumulación

5 Transformaciones geométricas

5.1 Traslaciones

En caso de las transformaciones geométricas tenemos que solucionar dos problemas distintos: Sean dadas dos listas L y L' de puntos. Si conocemos las parejas de puntos correspondientes, es decir para $p=(x,y) \in L$ y $p'=(x',y') \in L'$ la transformación $p \Rightarrow p'$ es válida, entonces podemos determinar la transformación con ayuda del método de los mínimos cuadrados.

Por ejemplo, una traslación es descrita por los dos ecuaciones

$$\mathbf{x}' = \mathbf{x} + \mathbf{c} \quad \mathbf{y} \quad \mathbf{y}' = \mathbf{y} + \mathbf{d}$$

Si hay muchas parejas de puntos correspondientes y si las coordenadas son perturbadas solamente por pequeños errores, entonces se tiene que solucionar el problema de la optimización

$$S(c, d) = \sum_{i=1}^N (x_i + c - x'_i)^2 + (y_i + d - y'_i)^2 \rightarrow \text{mínimo}$$

Con N como el número total de los parejas resulta

$$c = \frac{1}{N} \sum_{i=1}^N (x_i - x'_i)^2, \quad d = \frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2$$

Por otro lado, si conocemos la transformación \mathbf{T} , entonces podemos calcular los puntos nuevos $p_i^* = \mathbf{T}p_i$ y buscar aquel punto $q_{k_i} \in L'$ cuya distancia a p_i^* es un mínimo. De esta manera obtenemos los puntos correspondientes $q_{k_i} \leftrightarrow p_i$.

En casi todos los casos prácticos, la información disponible consta solamente en dos listas L y L' de puntos donde estas listas contienen también puntos sin un equivalente correspondiente.

Esta tarea del apareamiento de puntos (PPM, *point pattern matching*) se había investigado muchas veces (por ejemplo, Ranade y Rosenfeld, 1980; Cox y De Jager, 1993; Voss y Suesse, 2001). Aquí investigamos el problema PPM por el enfoque de la acumulación.

Sean dadas las dos listas de puntos L y L' . Entonces calculamos todas las diferencias

$$(c_{i,j}, d_{i,j}) = (x'_j - x_i, y'_j - y_i)$$

para $(x_i, y_i) \in L$, $(x'_j, y'_j) \in L'$ y acumulamos estas diferencias en el plano de las coordenadas c y d .

En la figura 13 vemos un ejemplo: Están dados 48 puntos (x,y) en cada una de las imágenes del tamaño 256^2 . No sabemos cuales de los puntos de la imagen izquierda y de la imagen derecha son puntos correspondientes (las coordenadas x y y son erróneas con un ruido de ± 1 pixel).

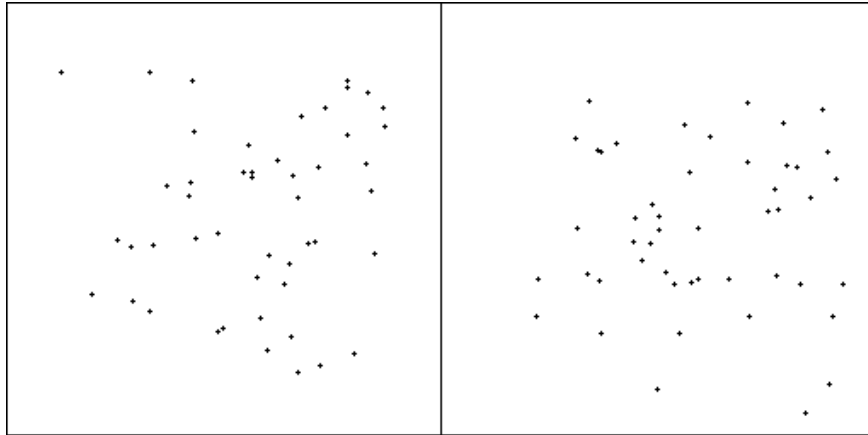


Figura 13: Imágenes con 8 puntos correspondientes

El método de RANSAC o de la acumulación necesita la calculación de $48^2 = 2304$ diferencias (c,d) . Estos resultados son distribuidos sobre todo el plano (c,d) , es decir en nuestro caso distribuidos arbitrariamente sobre una imagen (c,d) con 256^2 células (las diferencias pueden tomar valores entre -256 y $+256$). Pero las 8 diferencias (c,d) de los parejas de puntos correspondientes se concentran en sólo una célula. De esta manera obtenemos el resultado de la figura 14.

El pico de acumulación (figura 14, parte arriba a la izquierda) es muy destacado ya en esta imagen. Si acumulamos los valores de la calidad de transformación (es decir el número de los puntos muy cercanos después de la transformación), entonces el pico correcto se amplifica.

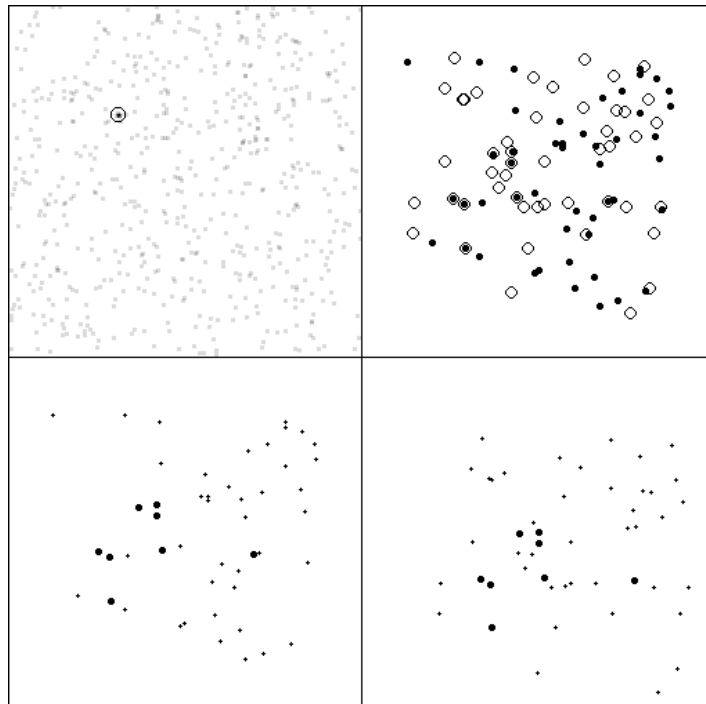


Figura 14: Pico de acumulación y el resultado de PPM

Para el pico correcto (c^*, d^*) obtenemos la transformación correcta y el resultado de la figura 14, parte arriba a la derecha. Las parejas de los puntos muy cercanos señalan los puntos correspondientes de las imágenes de la figura 13 (ver las partes bajas de la figura 14).

5.2 Transformaciones homotéticas

Los movimientos (traslación y rotación) son transformaciones un poco más complicadas que las traslaciones y son determinadas por los coeficientes c y d de la traslación y el ángulo φ de la rotación.

Las transformaciones homotéticas contienen también un factor f de la escala así a partir de las ecuaciones $x' = f \cdot (x \cos \varphi - y \sin \varphi) + c$, $y' = f \cdot (x \sin \varphi + y \cos \varphi) + d$ dos parejas de puntos correspondientes son suficientes para determinar los parámetros de la transformación:

$$\begin{aligned}x_1' &= f \cdot x_1 \cos \varphi - f \cdot y_1 \sin \varphi + c \\y_1' &= f \cdot x_1 \sin \varphi + f \cdot y_1 \cos \varphi + d \\x_2' &= f \cdot x_2 \cos \varphi - f \cdot y_2 \sin \varphi + c \\y_2' &= f \cdot x_2 \sin \varphi + f \cdot y_2 \cos \varphi + d\end{aligned}$$

Se pueden eliminar los parámetros de la traslación por una sustracción

$$\begin{aligned}x_1' - x_2' &= f \cdot (x_1 - x_2) \cos \varphi - f \cdot (y_1 - y_2) \sin \varphi \\y_1' - y_2' &= f \cdot (x_1 - x_2) \sin \varphi + f \cdot (y_1 - y_2) \cos \varphi\end{aligned}$$

y el factor de la escala por la ecuación

$$f = \sqrt{\frac{(x_1 - x_2)^2 + (y_1 - y_2)^2}{(x_1' - x_2')^2 + (y_1' - y_2')^2}} = \sqrt{\frac{l}{l'}}$$

Al final obtenemos el ángulo φ a partir de las relaciones

$$\begin{aligned}(x_1 - x_2)(y_1' - y_2') - (y_1 - y_2)(x_1' - x_2') &= ll' \sin \varphi \\(x_1 - x_2)(x_1' - x_2') - (y_1 - y_2)(y_1' - y_2') &= ll' \cos \varphi\end{aligned}$$

En la figura 15 se puede ver los resultados de la acumulación (proyecciones a los planos (f, φ) y (c, d)) a los 13 parejas de los puntos correspondientes.

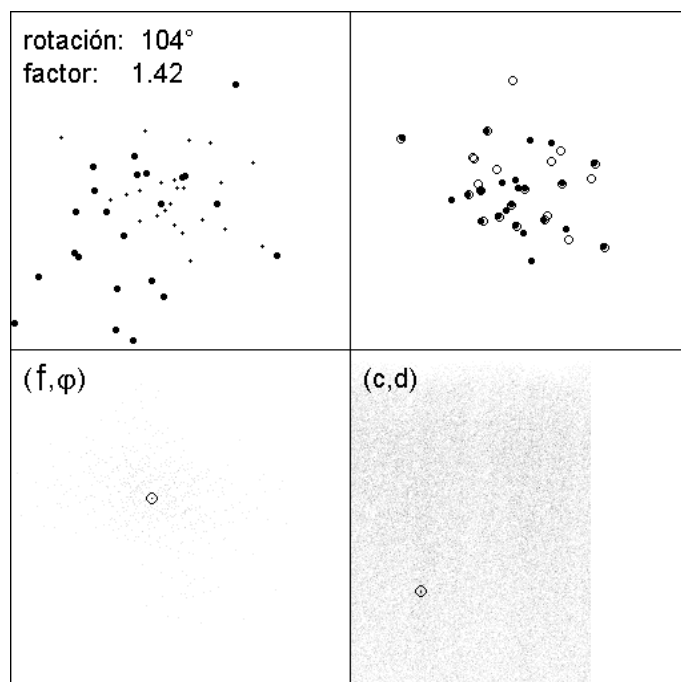


Figura 15: Puntos transformadas homotéticamente

5.3 Transformaciones afines

Una transformación T afín es definida por

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{10} \\ a_{21} & a_{22} & a_{20} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad \text{o} \quad \mathbf{x}' = \begin{pmatrix} \mathbf{B} & \mathbf{a}_0 \\ \mathbf{0}^T & 1 \end{pmatrix} \mathbf{x}$$

es decir, por 6 parámetros. La transformación T describe traslaciones, rotaciones, escalaciones no-isotrópicas y estiramientos (*shearings*). Se tienen que usar 3 parejas de puntos correspondientes para determinar los 6 parámetros:

$$x'_i = a_{10} + a_{11}x_i + a_{12}y_i \quad , \quad y'_i = a_{20} + a_{21}x_i + a_{22}y_i \quad \text{para } i \in \{1,2,3\}$$

Las diferencias entre estas ecuaciones proporcionan las fórmulas siguientes:

$$\begin{pmatrix} a_{11} \\ a_{12} \end{pmatrix} = \begin{pmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} x'_2 - x'_1 \\ x'_3 - x'_1 \end{pmatrix}$$

$$\begin{pmatrix} a_{21} \\ a_{22} \end{pmatrix} = \begin{pmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \end{pmatrix}^{-1} \cdot \begin{pmatrix} y'_2 - y'_1 \\ y'_3 - y'_1 \end{pmatrix}$$

$$a_{10} = x_1' - a_{11}x_1 - a_{12}y_1$$

$$a_{20} = y_1' - a_{21}x_1 - a_{22}y_1$$

En lugar de la descripción de una transformación afín por los seis parámetros algebraicos a_{ik} podemos usar también seis parámetros geométricos descomponiendo la matriz \mathbf{T} como un producto $\mathbf{T}=\mathbf{HAD}$ con una transformación homogénea \mathbf{H} (rotación y escalación isotrópica con respecto al origen), una traslación o desplazamiento \mathbf{D} con dos parámetros c,d y una “transformación afín pura” \mathbf{A} que describe la escalación no-isotrópica por un parámetro s y el estiramiento por un parámetro t :

$$\mathbf{T} = \begin{pmatrix} f \cos \varphi & -f \sin \varphi & 0 \\ f \sin \varphi & f \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} s & st & 0 \\ 0 & 1/s & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & c \\ 0 & 1 & d \\ 0 & 0 & 1 \end{pmatrix}$$

En la mayoría de los casos prácticos podemos asumir que $t \approx 0$ y $s \approx 1$ así que el pico más destacado en el espacio 6-dimensional de la acumulación está localizado cerca del punto $(f_{\text{opt}}, \varphi_{\text{opt}}, 1, 0, c_{\text{opt}}, d_{\text{opt}})$.

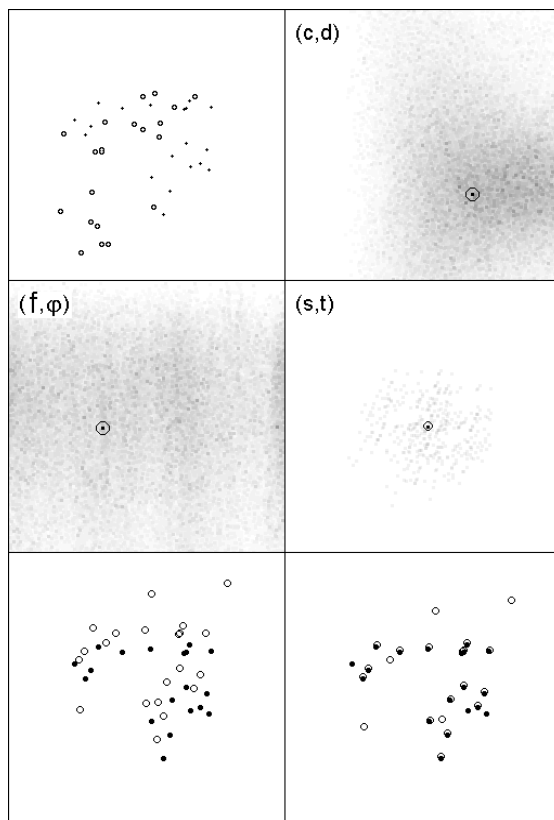


Figura 16: Determinación de una transformación afín

En la figura 16 mostramos un ejemplo con las proyecciones de la acumulación a los planos (f,φ) , (s,t) y (c,d) .

La imagen arriba a la izquierda muestra la tarea: se tiene que detectar las parejas de puntos correspondientes y la transformación afín que mapea optimamente el primer conjunto de los puntos al segundo conjunto.

Las dos imágenes de abajo contienen el resultado sin traslación (a la izquierda) y el resultado completo de la transformación determinada (a la derecha). Se puede ver que las 15 puntos correspondientes son hallados más o menos exactamente y que también 3 o 5 puntos sin compañeros son determinados.

6 Detección de desplazamientos

6.1 Convoluciones y operadores lineales

Un operador lineal que transforma funciones discretas periódicas $F=(F_0,F_1,\dots,F_{N-1})$ por $G=LF$ con $F,G \in F$ es caracterizado por

$$\mathbf{L}(a \cdot F + b \cdot G) = \mathbf{L}(a \cdot F) + \mathbf{L}(b \cdot G) = a \cdot \mathbf{L}F + b \cdot \mathbf{L}G$$

con a,b como números reales y N como la longitud del período. Si una función F es desarrollada como una suma sobre funciones básicas $B^{(n)}$ entonces tenemos

$$\mathbf{L}F = \mathbf{L} \left(\sum_{n=0}^{N-1} f_n B^{(n)} \right) = \sum_{n=0}^{N-1} f_n \mathbf{L}(B^{(n)}).$$

con los coeficientes f_n . Las funciones $\mathbf{L}(B^{(n)})$ podemos escribir como

$$\mathbf{L}B^{(n)} = \sum_{m=0}^{N-1} \lambda_m^{(n)} B^{(m)}$$

Por eso obtenemos que la aplicación de un operador lineal \mathbf{L} a una función F es lo mismo que la multiplicación de esta función F con una “matriz circulante” \mathbf{L} con los coeficientes $(\mathbf{L})_{n,m} = L_{n,m} = \lambda_m^{(n)}$.

Una convolución $F*G$ entre dos funciones es definida por la fórmula

$$(F*G)_n = \sum_{k=0}^{N-1} f_{n-k} g_k$$

y tiene las propiedades

$$F*(G + H) = F*G + F*H$$

$$F*G = G*F \quad , \quad F*(G*H) = (F*G)*H$$

Se llama un operador lineal C_F como *operador de convolución* si hay una función discreta periódica de tal manera que

$$C_F G = F * G$$

para todas las funciones $G \in F$.

El desplazamiento de una función F por k lugares proporciona una función nueva F^k con los coeficientes

$$f_n^k = f_{n-k}$$

Se llama un operador L como “operador shift invariant” si

$$L F^k = (L F)^k$$

Cada operador de convolución C_F es *shift invariant* por que podemos comprobar que

$$\begin{aligned} (C_F G^k)_n &= (F * G^k)_n = \sum_{m=0}^{N-1} f_m g_{n-m}^k \\ &= \sum_{m=0}^{N-1} f_m g_{n-k-m} = (F * G)_{n-k} = ((F * G)^k)_n = ((C_F G)^k)_n \end{aligned}$$

Operadores lineales que son *shift invariant* se llaman “operador LSI”, y por eso cada operador de convolución es un operador LSI.

Si desarrollamos una función F en una suma sobre las funciones básicas $E^{(n)}$ con coeficientes $E_k^{(n)} = \delta_{k-n}$ (con $\delta_0=1$ y $\delta_m=0$ para $m \neq 0$) entonces obtenemos a partir de la relación

$$(E^{(k)})_n = (E^{(0)})_{n-k}$$

el resultado

$$(L G)_n = \sum_{k=0}^{N-1} g_k (L E^{(k)})_n = \sum_{k=0}^{N-1} g_k ((L E^{(0)})^k)_n = \sum_{k=0}^{N-1} g_k (L E^{(0)})_{n-k} = (G * L E^{(0)})_n$$

La función $L E^{(0)}$ se llama “*discrete point spread function*” o PSF por que en el campo del procesamiento de imágenes los operadores LSI las más veces son operadores con coeficientes muy pequeños (o más aún ceros) afuera de un rango restringido de los índices.

La convolución $F * G$ de dos funciones $F, G \in F$ es equivalente a la aplicación de un operador LSI C_F a la función G y también equivalente a la multiplicación del vector $G = (G_0, G_1, \dots, G_{N-1})$ por una matriz circulante C_F :

$$\mathbf{L}G = \mathbf{C}_F G = C_F G = \begin{pmatrix} f_0 & f_{N-1} & f_{N-2} & \cdot & \cdot & \cdot & f_2 & f_1 \\ f_1 & f_0 & f_{N-1} & \cdot & \cdot & \cdot & f_3 & f_2 \\ f_2 & f_1 & f_0 & \cdot & \cdot & \cdot & f_4 & f_3 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ f_{N-2} & f_{N-3} & f_{N-4} & \cdot & \cdot & \cdot & f_0 & f_{N-1} \\ f_{N-1} & f_{N-2} & f_{N-3} & \cdot & \cdot & \cdot & f_1 & f_0 \end{pmatrix} \begin{pmatrix} G_0 \\ G_1 \\ G_2 \\ \cdot \\ \cdot \\ G_{N-2} \\ G_{N-1} \end{pmatrix}$$

6.2 Inversión de convoluciones

Ahora investigamos el problema general de la solución de la ecuación $F * X = E^{(0)}$, es decir la determinación de la inversa de una función (en el sentido de las convoluciones).

No es necesario desarrollar una función en una suma sobre las funciones unitarias $E^{(n)}$ sino usamos funciones básicas generales $B^{(n)}$. Entonces escribimos una convolución como sigue:

$$F * G = \sum_{k=0}^{N-1} f_k B_k * \sum_{l=0}^{N-1} g_l B_l = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f_k g_l B_k * B_l$$

La convolución de dos funciones básicas B_k y B_l es

$$B_k * B_l = \sum_{m=0}^{N-1} \gamma_{k,l}^{(m)} B_m$$

por que podemos desarrollar también la nueva función $B_k * B_l$ en una suma sobre las funciones básicas. Entonces el resultado es

$$F * G = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} f_k g_l B_k * B_l = \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} \sum_{m=0}^{N-1} f_k g_l \gamma_{k,l}^{(m)} B_m$$

Llamamos los coeficientes γ *constantas de la estructura*. Ellos dependen de las funciones básicas elegidas. Para una *base canónica* estos coeficientes tienen los valores siguientes:

$$\gamma_{k,l}^m = \begin{cases} 1 & \text{para } k=l=m \\ 0 & \text{para todos otros casos} \end{cases}$$

En este caso obtenemos

$$F * G = \sum_{k=0}^{N-1} f_k g_k B_k \quad \circ \quad (F * G)_k = f_k g_k$$

es decir, que los coeficientes del resultado son definidos simplemente por el producto de los coeficientes de las dos funciones originales. Funciones básicas de este tipo tienen la estructura

$$B_n = \frac{1}{N} (\mathbf{a}^0, \mathbf{a}^n, \mathbf{a}^{2n}, \dots, \mathbf{a}^{(N-1)n})$$

con una variable algebraica \mathbf{a} que tiene que satisfacer las condiciones (ver Voss 1993)

$$\mathbf{a}^N = 1 \quad , \quad N^{-1} \text{ existe}$$

$$(\mathbf{a}^k - 1)^{-1} \text{ existe para } k \in \{1, 2, \dots, N-1\}$$

Un ejemplo de estas condiciones es $\mathbf{a} = e^{2\pi i/N}$, es decir la base de la transformación de Fourier.

La inversa de una función F con respecto a la convolución $F * X = E^{(0)}$ está dada por los coeficientes ξ_k de la función transformada de la función desconocida X :

$$\xi_k = \frac{\epsilon_k^{(0)}}{\varphi_k} = \frac{1}{\varphi_k} \quad , \quad X_k = \frac{1}{N} \sum_{l=0}^{N-1} \xi_l \cdot e^{2\pi i kl/N} = \frac{1}{N} \sum_{l=0}^{N-1} \frac{e^{2\pi i kl/N}}{\varphi_l}$$

Pero hay problemas si cualquiera de los coeficientes φ_k es cero. Además estaría bien si una inversa de una función restringida localmente de nuevo estuviera restringida localmente. Normalmente no hay una solución de estos deseos (por ejemplo en el caso de $\varphi_k=0$). Por eso buscamos una solución aproximada X con la condición $F * X \approx E^{(0)}$.

Una solución X de este tipo es caracterizado por la condición que el producto escalar

$$S = \langle F * X - E^{(0)}, F * X - E^{(0)} \rangle$$

obtiene un valor mínimo. Con la suposición de que X es una solución óptima, una pequeña variación ΔX no cambia el valor S hasta expresiones del grado dos, es decir que $S + \Delta S = S$ y $\Delta S = 0$. Por eso obtenemos

$$\Delta S = \langle F * X - E^{(0)}, F * \Delta X \rangle + \langle F * \Delta X, F * X - E^{(0)} \rangle + \langle \Delta X, \Delta X \rangle .$$

Si abandonamos la expresión $\langle \Delta X, \Delta X \rangle$ y consideramos la fórmula

$$\langle F * G, H \rangle * = \langle G, F^{\leftrightarrow} * H \rangle$$

para la combinación de convoluciones y productos escalares, entonces obtenemos

$$\Delta S = 2 \langle (F * X - E^{(0)}) * F^{\leftrightarrow}, \Delta X \rangle = 0$$

con F^{\leftrightarrow} como la función reflejada $(F^{\leftrightarrow})_k = F_{-k}$. Porque las variaciones ΔX_k son arbitrarias podemos decir que los coeficientes del lado izquierdo del producto escalar ΔS tienen que ser cero:

$$\left((F * X - E^{(0)}) * F^{\leftrightarrow} \right)_q = 0 \quad \text{for } q \in Q.$$

Aquí Q es el conjunto de todos los puntos para que los coeficientes X_q pueden ser distintos de cero, es decir el dominio acotado de la solución X . Finalmente obtenemos

$$\left(F * F^{\leftrightarrow} * X \right)_q = \left(F^{\leftrightarrow} \right)_q \quad \text{for } q \in Q$$

y por eso

$$F * F^{\leftrightarrow} * X = F^{\leftrightarrow}$$

con X como la función pseudoinversa que buscábamos.

Ahora podemos determinar la función pseudoinversa X por la transformación Fourier por sus coeficientes ξ_k :

$$\xi_k = \frac{\varphi_k^*}{\varphi_k \varphi_k^* + \beta} \quad (\text{A})$$

Con $\beta=0$ tenemos la solución “exacta” de la ecuación $F * F^{\leftrightarrow} * X = F^{\leftrightarrow}$. Pero con $\beta > 0$ no hay problemas si cualquiera de los coeficientes φ_k es cero: El producto $\varphi_k \varphi_k^*$ nunca es negativo y por eso el denominador siempre es positivo si elegimos un pequeño valor positivo para el parámetro β .

6.3 El método SDR

Es más simple si escribimos todas las fórmulas para solo una dimensión. Sean f y f' dos imágenes, δ la función del impulso unitario y \mathbf{S}_d el operador del desplazamiento por la distancia d . Entonces obtenemos la fórmula básica

$$f' = \mathbf{S}_d f = (\mathbf{S}_d f) * \delta = f * \mathbf{S}_d = f * \delta_d$$

Para los coeficientes de Fourier obtenemos el siguiente teorema ($\alpha_k(f)$ son los coeficientes discretos de Fourier de la función f):

$$\alpha_k(f') = \alpha_k(f) \cdot e^{2\pi i \frac{kd}{N}}$$

Este teorema del desplazamiento (*shift theorem*) es la base de todos los métodos basados en las fases (se por ejemplo Fleet y Jepson, 1990; Fleet y Jepson, 1991; Weng 1993; Xiong y Shafer 1997). Pero el teorema del desplazamiento usa solamente un valor de la frecuencia (un valor de k). Por eso el efecto integral de la convolución (empleo de muchas frecuencias) se pierde.

Aquí presentamos un método nuevo que tiene en cuenta no sólo el desplazamiento de las imágenes sino también cambios de la función de gris y ruido (primeros resultados se han descrito en Voss et al., 1998).

En cierto sentido el método SDR es similar al método clásico de “matched filtering random signal processing technique” (ver Pratt, 1991). Pero el método SDR no determina un máximo de la relación señal/ruido, sino restaura el impulso δ desplazado. Porque determinamos el desplazamiento (*shift*) por un método de la restauración, llamamos el método SDR (*shift detection by restoration*).

Porque las imágenes f y f' contienen perturbaciones y ruido, completamos la ecuación básica por una imagen p que representa todos estos efectos adicionales:

$$f' = f * \delta_d + p$$

Si investigamos la solución de la tarea $S=|p|^2 \rightarrow \text{mínimo}$ con $p = f' - f * x$ (con x como una imagen aproximada del impulso δ desplazada), entonces obtenemos otra vez como en la sección 6.2 una ecuación

$$\Delta S = 2 \langle (f * x - f') * f'^*, \Delta x \rangle = 0$$

o

$$f * f'^* * x = f' * f'^*$$

La ecuación siguiente corresponde a la ecuación (A) de la sección 6.2 (también aquí es introducido un término β para tener en cuenta las perturbaciones):

$$\xi_k = \frac{\varphi_k' \varphi_k^*}{\varphi_k \varphi_k^* + \beta} \quad (B)$$

El método SDR logra resultados buenos también en el caso de perturbaciones grandes como en la figura 17 (desplazamiento grande, rotación por 5 grados, cambio de la función de gris, ruido). La imagen arriba derecha muestra la simple “deconvolución pura” con $\beta=0$ y $\xi_k = \varphi_k' / \varphi_k$.

Las imágenes en la mitad de figura 17 se determinaron con valores $\beta=10^2$, $\beta=10^3$ y $\beta=10^4$ (los círculos pequeños señalan los picos en las imágenes x). El valor apropiado de β depende de la normalización de los coeficientes Fourier.

La tercera línea de la figura 17 muestra las imágenes de la SDR para $\beta=10^6$, de la “correlación cruzada pura” ($\beta \rightarrow \infty$) y de la “correlación de fase”.

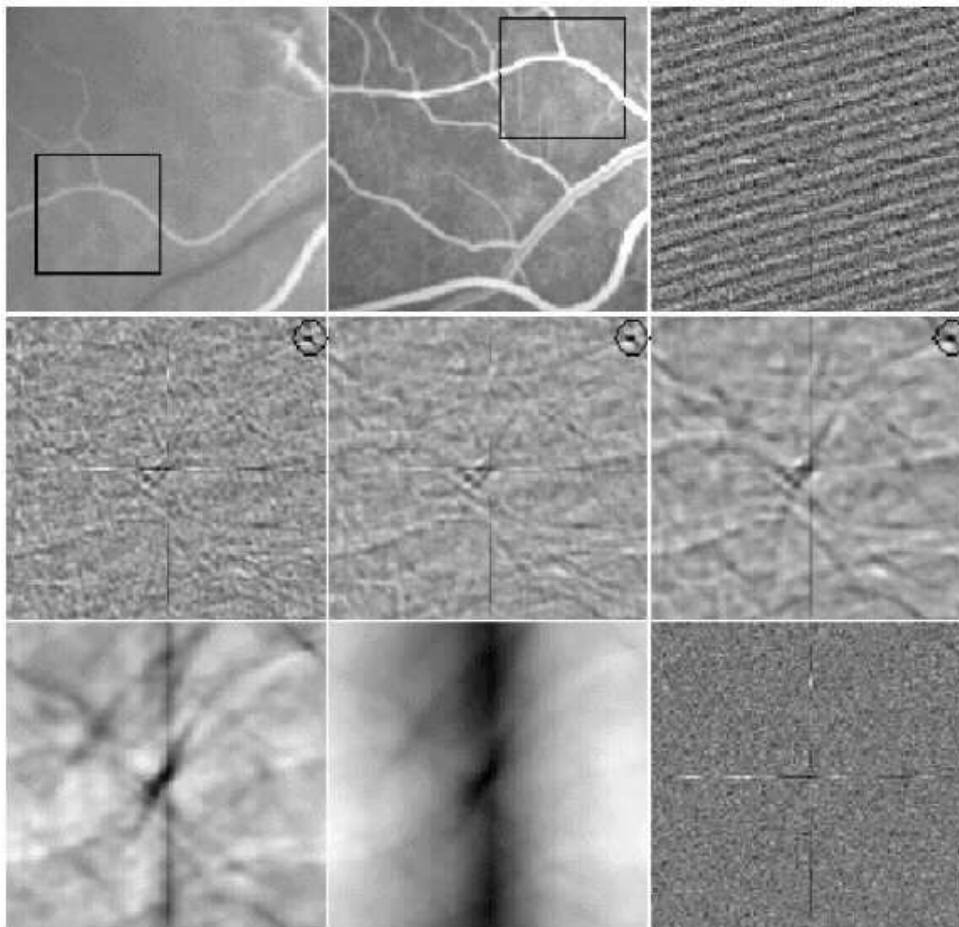


Figura 17: Dos imágenes (con desplazamiento, rotación y cambio de los valores de gris) y la detección del desplazamiento por el método SDR

En muchos casos no hay solo una traslación sino transformaciones más complejas entre dos imágenes:

$$x' = \xi(x,y) \quad , \quad y' = \eta(x,y)$$

Pequeñas alteraciones en los coordenadas (x,y) resultan en pequeños cambios de x' y y' . En caso de que no hay una rotación por un ángulo más grande de unos grados y no hay un factor de escalamiento muy diferente de 1.0, cada punto y su entorno son desplazados por una traslación aproximada. Bajo estas condiciones es posible determinar el desplazamiento entre dos imágenes en ventanas pequeñas (por ejemplo con un tamaño de 32×32 píxeles):

- imagen 1 del tamaño $n \times n$ con ventana del tamaño $k \times k$ al lugar (x_v, y_v)
- imagen 2 del tamaño $n \times n$ con ventana del tamaño $k \times k$ al lugar (x_v, y_v)

Para cada ventana hay generalmente otro desplazamiento, y por eso obtenemos un campo de desplazamientos (ver figura 18).

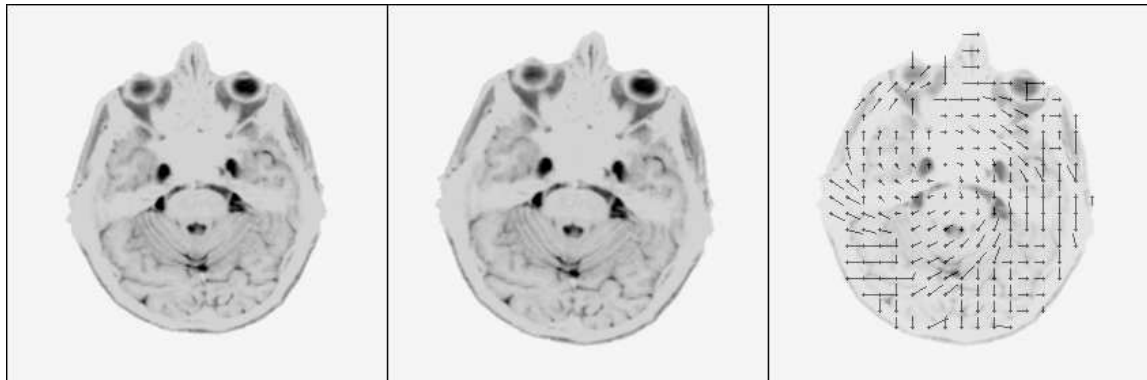


Figura 18: Dos imágenes (con rotación, traslación y escalamiento) y el campo de desplazamientos (dependiente de lugar)

Si determinamos el desplazamiento $(\Delta x, \Delta y)$ de las ventanas por el método SDR, entonces obtenemos una lista de puntos correspondientes en las dos imágenes con coordenadas (x, y) y (u, v) :

$$\begin{array}{lcl}
 \text{imagen 1} & (x_1, y_1) \Leftrightarrow & (u_1, v_1) \quad \text{imagen 2} \\
 & (x_2, y_2) \Leftrightarrow & (u_2, v_2) \\
 & \dots\dots\dots & \\
 & (x_m, y_m) \Leftrightarrow & (u_m, v_m)
 \end{array}$$

Ahora se puede calcular la transformación supuesta (traslación, movimiento, transformación afín etc.) porque las correspondencias puntuales son conocidas.

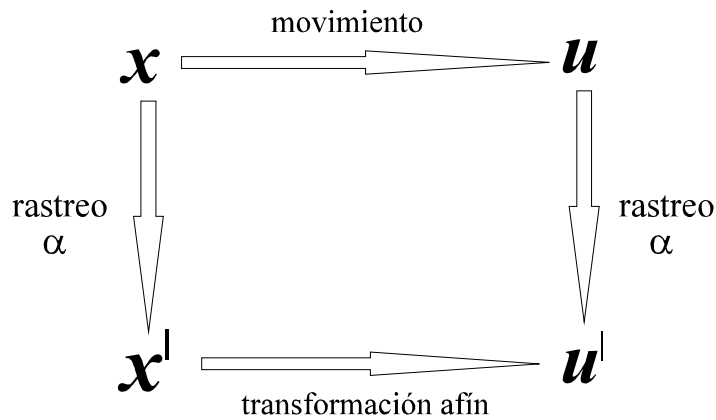
6.4 Investigaciones experimentales

Como aplicación práctica queremos mostrar un ejemplo bidimensional que ya es investigado en el caso tridimensional por (Botello y Marroquín, 2003):

“El problema de registro en imágenes consiste en encontrar la transformación geométrica que ponga dos imágenes dadas en la mejor correspondencia posible”.

En la práctica de la tomografía computerizada (imágenes tresdimensionales del tipo CT, MRT, MRI) los coordenadas x y y de un lado y z de al otro lado tienen resoluciones distintas (hasta factores de 3). Por eso un movimiento de una cabeza entre dos captaciones (o rastreos) no corresponde a un movimiento de las imágenes captadas.

Cada una de los dos objetos movidos (por ejemplo de las cabezas) es deformado por una compresión en la dirección del coordenada z :



Hasta una traslación, el movimiento está realizado por una rotación (matriz de rotación \mathbf{R}). La transformación \mathbf{T} entre \mathbf{x}' y \mathbf{u}' es una transformación afín especial:

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \mathbf{x}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{C}\mathbf{x}$$

$$\mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \Rightarrow \mathbf{u} = \begin{pmatrix} u \\ v \\ w \end{pmatrix} = \mathbf{R} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{R}\mathbf{x}$$

y por eso

$$\mathbf{u} \Rightarrow \mathbf{u}' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \alpha \end{pmatrix} \mathbf{u} = \mathbf{C} \cdot \mathbf{R}\mathbf{x} = \mathbf{C} \cdot \mathbf{R}\mathbf{C}^{-1} \mathbf{x}' = \mathbf{T}\mathbf{x}'$$

En la figura 19 se puede ver la efectividad del método SDR (fue realizada una compresión de un factor 1.6 en la dirección del ordenada). El campo de desplazamiento fue determinado por ventanas del tamaño 32×32 , y a partir de estas desplazamientos una transformación afín fue calculado. Las imágenes de diferencia muestran, que la transformación fue determinada casi correctamente.

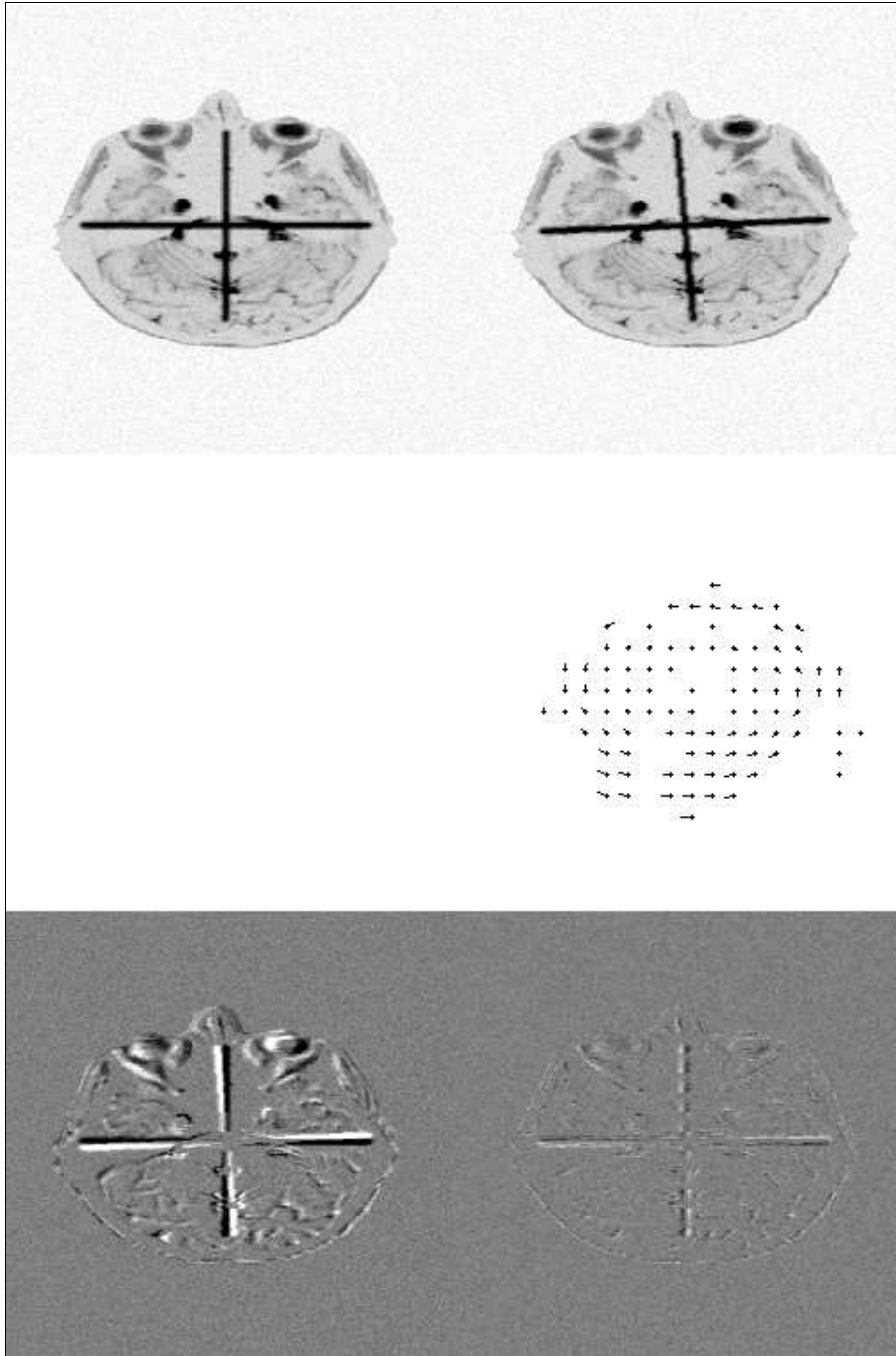


Figura 19: Dos imágenes bidimensionales (con rotación, traslación, compresión y ruido), campo de desplazamiento, diferencias originales y diferencias después el registro

7 Conclusiones

Hemos desarrollado un enfoque único para la detección de figuras geométricas simples como líneas rectas, círculos y elipses por imágenes o listas de puntos. La porción de los puntos perturbados puede ser bastante grande, por ejemplo considerablemente mayor de 50%.

Este enfoque se funda en la acumulación de los parámetros determinados a partir un conjunto mínimo de puntos (dos puntos para líneas, tres puntos para círculos etc.). En este sentido el enfoque es similar al método RANSAC (Fischler y Bolles, 1981).

Pero al contrario de RANSAC no calculamos solamente un resultado único y suficiente sino un conjunto de resultados (eligidos arbitrariamente) y determinamos el resultado óptimo por la acumulación de los resultados elementales. Se pueden mejorar los resultados por la acumulación de valores de calidad (número de los puntos a lo largo de la línea etc.).

La idea de la acumulación se puede extender también a la tarea de la determinación de transformaciones geométricas. Para parejas de puntos (o parejas de parejas de puntos, parejas de tripletas de puntos) podemos determinar la transformación (traslación, transformación homotética, transformación afín) y las parejas de puntos correspondientes. La porción de los puntos de ruido (puntos sin equivalentes correspondientes) supera también aquí el porcentaje de 50%.

En el parte 6 el método SDR (*shift detection by restoration*) es descrito. A partir de este método se puede determinar puntos correspondientes en dos imágenes y después por estos parejas de puntos se determina la transformación entre ambos imágenes.

Agradecemos a Dr. Humberto Sossa (CIC, México D.F.) por su apoyo en preparación y realización del trabajo. Igualmente queremos agradecer a Dr.J.L.Marroquín (CIMAT) por su ayuda continua durante la estancia de K.Voss en Guanajuato (de enero hasta junio 2004).

8 Referencias

- Anger,B., Portenier,C.** "Radon Integrals", Birkhäuser, Boston MA 1992
- Ballard, D.H.** "Generalizing the Hough transform to detect arbitrary shapes", Pattern Recognition 13 (1981) 111-122
- Cao,X., Deravi,F.** " An efficient method for multipli-circle detection", Proc.IEEE Conf. on Computer Vision 1990, pp. 744-746
- Cox,G.S., DeJager,G.** "A survey of point pattern matching techniques and a new approach to point pattern recognition", In: South Afracan Symposium on Communications and Signal Pr5ocessing 1993, pp. 243-248
- Davies, E.R.** "Machine Vision", Academic Press 1990, chapter 9
- Deans,S.R.** "The Radon Transform and Some of Its Applications", Wiley, New York 1983
- Fischler,M.A., Bolles,R.C.** "Random sample consensus - a paradigm for model fitting with applications to image analysis and automated cartography", Comm. of the ACM 24 (1981) 381-395

Fleet D.J., Jepson A.D. "Computation of component image velocity from local phase information ", IJCV 5 (1990) 77-104

Fleet D.J., Jepson A.D., Jenkin M.R.M. "Phase-based disparity measurement ", CVGIP Image Understanding 53 (1991) 198-210

Gonzalez,R.C., Woods,R.E. "Digital Image Processing", Addison-Wesley Comp. 1993

Goshtasby,A., Stockmann,G.C. "Point pattern matching using convex hull edges", IEEE Trans. SMC-15 (1985) 631-637

Guil,N., Zapata,E.L. "Lower order circle and ellipse Hough transform", Pattern Recognition 30 (1997) 1729-1744

Hough,P.V.C. "Methods and means for recognizing complex patterns", US patent 3069654 (1962)

Kimme,C., Ballard,D.H., Sklansky,J. "Finding circles by an array of accumulators", CACM 18 (1975) 120-122

Li,H., Lavin,M.A., LeMaster, R.J. "Fast Hough transform - a hierarchical approach", CVGIP 36 (1986) 139-161

Botello,S., Marroquín,J.L. "El método de los elementos finitos en el registro de imágenes de resonancia magnética", Comunicación Técnica I-03-08/7-04-2003, CIMAT Guanajuato 2003

Pratt,W.K. "Digital Image Processing", Wiley&Sons, New York 1991

Radon,J. "Über die Bestimmung von Funktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten", Berichte und Verhandlungen der Sächsischen Akademie der Wissenschaften, Math./Nat. Klasse 69 (1917) 262-277

Ranade,S., Rosenfeld,A. "Point pattern matching by relaxation", Pattern Recognition 12 (1980) 269-275

Rosin,P. "A note on least squares fitting of ellipses", Pattern Recognition Letters 14 (1993)

Sonka,M., Hlavac,V., Boyle,R. "Image Processing, Analysis and Machine Vision", Chapman&Hall 1993

Voss,K., Suesse,H. "Adaptive Modelle und Invarianten für zweidimensionale Bilder", Shaker Verlag, Aachen 1995, pp. 12-13

Voss,K., Suesse,H. "Apareamiento de puntos y optimización lineal", Comunicaciones del CIMAT/Guanajuato/ México, ComunicaciónTécnica 1-01-06/05-06-2001

Voss,K., Suesse,H. "Affine point pattern matching", Proc. 23rd DAGM Sympos. 2001, München, Sept.2001, Lecture Notes in Computer Science, Springer 2001, pp.155-162

Voss,K., Marroquin.J.L., Gutiérrez, S., Suesse,H. "Fundamentos y Nuevos Métodos para la Reconstrucción Tridimensional", cap. 7 , secc. 7.4, ver en Internet (1999): http://www.cimat.mx/proy_comp/libror3d/libror3d.html

Voss,K., Marroquin.J.L., Gutiérrez, S., Suesse,H. "Analysis de Imágenes de Objetos Tridimensionales", CIC México City 2002, secc.7.4, en prensa

Biografías de los autores



Klaus Voss recibió el grado de doctor en física teórica en 1965 en la Universidad Técnica de Dresden (Alemania), y el grado de doctor habilitado (Dr.habil.) en 1969 en la misma universidad. El ha trabajado como profesor asistente en el Depto. de Física y más tarde en el Depto. de Informática de la Universidad Técnica de Dresden. Desde 1971 ha trabajado en el área del procesamiento digital de imágenes, por ejemplo durante muchos años en el Depto. de Patología del Hospital "Charité" de la Universidad "Humboldt" de Berlin. En 1986 el fue empleado como catedrático y jefe de un nuevo grupo "Procesamiento de Imágenes" en la Universidad "Friedrich Schiller" de Jena (Alemania), y tiene hasta la fecha esta posición. Dr.Voss ha publicado más de 260 artículos científicos y es autor/co-autor de 7 libros sobre el procesamiento digital de imágenes. Su intereses estan en la teoría del procesamiento de imágenes, software interactivo, invariantes, reconocimiento de formas en imágenes y reconstrucción tridimensional.



Herbert Suesse recibió el grado de doctor en matemáticas (análisis numérica) en 1973 en la Universidad Técnica de Merseburg (Alemania) y el grado de doctor habilitado (Dr.habil.) en el campo de la optimización matemática en 1981 en la misma universidad. El ha trabajado como profesor asistente en el Depto. de Matemáticas de esta universidad. Desde 1980 ha trabajado en el Centro de Computación en Gera (Alemania). En 1987 el fue empleado como profesor asistente en el Depto. de Matemáticas y Ciencias Computacionales de la Universidad "Friedrich Schiller" de Jena, y tiene hasta la fecha esta posición. Los intereses principales de Dr. Suesse son la reconstrucción tridimensional, la teoría y la aplicación de los invariantes así como los métodos estadísticas en el campo de procesamiento de imágenes. Herbert Suesse ha publicado más que 60 artículos científicos y tres libros sobre problemas numéricas en la optimización, el procesamiento de imágenes y reconstrucción tridimensional.



Wolfgang Ortmann estudió en el departamento de física de la universidad "Friedrich Schiller" en Jena (Alemania) y trabajó unos años en este departamento (área principal de sus investigaciones: "superconductive sensors"). Desde 1987 el fue miembro del departamento de matemáticas de la universidad Jena. Enseñó y investigó en el campo de procesamiento de imágenes. El recibió el grado de doctor en ciencias computacionales (Dr.Ing.) en 1990 en el departamento de matemáticas de la universidad de Jena (Alemania). Después sus intereses principales fueron la desarrollo de software para procesamiento de imágenes y el tratamiento de imágenes basando en señales. Trabajó desde muchos años como maestro en la enseñanza universitaria básica (procesamiento de imágenes, fundamentos de informática). En los últimos años se dedicó a la visión robótica y la restauración de imágenes. Wolfgang Ortmann ha publicado más que 20 artículos científicos.