

# Rectificación Monocular de Imágenes

## *Monocular Rectification of Images*

Klaus Voss and Christian Bräuer-Burchardt

Universidad Friedrich Schiller, Depto. de Matemáticas y Ciencias Computacionales  
D-07740 Jena, Alemania, e-mail: nkvl@uni-jena.de

*Artículo recibido el 03 de marzo de 2000; aceptado el 06 de diciembre de 2000*

### Resumen

Presentamos un método nuevo para la rectificación de imágenes arquitectónicas que son captadas a partir de una lente denominada "ojo de pez" (fish-eye view). En muchos casos prácticos ni puntos de referencia ni imágenes estereográficas son disponibles, sino se tiene solamente imágenes monoculares. En este trabajo utilizamos propiedades típicas de los objetos mapeados como rectilínea, paralelidad, ortogonalidad, simetría y repetición de detalles.

El método consiste en dos pasos. El primer paso es la determinación y la corrección de la distorsión radial. Aquí usamos solamente informaciones sobre líneas curvadas en la imagen las cuales corresponden a líneas rectas en la realidad. El resultado del primer paso es una imagen sin distorsiones radiales, en la cual hay solamente distorsiones proyectivas (como en el caso de una cámara ideal).

El segundo paso es la corrección de las distorsiones proyectivas causadas por la proyección central del espacio tridimensional a una imagen bidimensional. Para lograr este objetivo usamos las relaciones entre puntos de fuga, el punto principal y la distancia focal de la cámara. En caso de que hay dos puntos de fuga para cada fachada, podemos determinar una transformación proyectiva especial que mapea la imagen de la fachada en una imagen rectificadas donde esta fachada aparece como una reproducción a escala.

Presentamos unos ejemplos para mostrar la eficacia del método. Nuestro enfoque es un método efectivo para rectificar imágenes monoculares de objetos arquitectónicos como primera solución – también en el caso que faltan la descripción y los datos de la cámara usada.

**Palabras Clave:** Rectificación de imágenes, Arquitectura, Fachadas, Puntos de fuga, Distorsión radial, Mejoramiento de imágenes, Transformaciones geométricas.

### Abstract

A new method for the rectification of architectonic images that are captured from a fish-eye view is presented. In many practical cases, neither reference nor stereographic points are available, instead it only has monocular images. This work uses typical properties of the mapped objects such as rectilinearity, parallelism, orthogonality, symmetry and repetition of details.

The method consists of two steps. The first step is the determination and the correction of the radial distortion. Here we only use information on curved lines on the image which corresponds to straight lines in reality. The result of the first step is an image without radial distortions, in which there are only projective distortions (as in the case of an ideal camera).

The second step is the correction of the projective distortions caused by the central projection of the three dimensional space to a bidimensional image. To reach this objective we used the relations between the vanishing points, the main point and the focal distance of the camera. In case that there are two vanishing points for each facade, we can determine a special projective transformation that maps the image of the facade in a rectified image where this facade appears as a scaled reproduction.

Examples are presented to show the method efficiency. Our approach is an effective method to rectify monocular images of architectonic objects as a first solution – also in the case where the description and data of the camera used is missing.

**Keywords:** Rectification of Images, Architecture, Facades, Vanishing Points, Radial Distortion, Improvement of Images, Geometric Transformations.

## 1 Introducción

En muchos campos de la visión computacional se deben lograr mediciones cuantitativas, por ejemplo hay muchas tareas de la arquitectura y de la generación de modelos CAD (computer aided design - diseño asistido por computadora) que se deben solucionar a base de las fotografías de los objetos (edificios, casas, puentes etc., ver Albers, 1995). Estas fotografías contienen distorsiones lineales (distorsiones proyectivas causadas por la proyección central del espacio real tridimensional a la imagen bidimensional) y distorsiones no lineales (causadas por el uso de cámaras de tipo *super-wide-angle* y cámaras de tipo *fish-eye*).

Además muchas veces no se utiliza cámaras estereográficas así que es necesario realizar una reconstrucción monocular. Por otro lado, cámaras con un objetivo gran angular y cámaras de tipo fish-eye tienen unas ventajas ya que se puede mapear espacios grandes en solo una fotografía y no se necesita combinar muchas imágenes separadas a una imagen compuesta. La eliminación de la distorsión no se realiza generalmente por una calibración de la cámara.

Hay muchas publicaciones sobre la calibración (Brown, 1971; Devernay, 1995; Nomura, 1992; Shih *et al.*, 1995; Shah, 1996; Tsai, 1986; Weng *et al.*, 1992). Métodos especiales (minimización no lineal directa, soluciones cerradas, soluciones en dos pasos) son discutidos en Weng *et al.*, 1992. Shah y Aggarwal (1996) describen una calibración para cámaras fish-eye con ayuda de un "campo de calibración" (un plano con puntos y líneas rectas). Sawhney (1999) y Stein (1997) presentan métodos que trabajan a partir de correspondencias entre los puntos en imágenes distintas.

Por lo general, una calibración precisa necesita gastos y esfuerzos grandes, y muchas veces una calibración no es posible si se tiene solamente las imágenes sin conocimientos sobre la cámara usada.

Aquí queremos investigar un enfoque general para la eliminación de la distorsión radial y la rectificación de las imágenes distorsionadas proyectivamente si hay sólo una imagen sin conocimientos sobre la cámara usada (ver por ejemplo figura 1.1).

Pero entonces necesitamos informaciones a priori sobre los objetos mapeados. En caso de los objetos arquitectónicos estas informaciones son conocimientos sobre la linealidad, el paralelismo, la ortogonalidad y la simetría de los objetos (en unas circunstancias también la repetición de algunos detalles del objeto). Como modelo usamos una cámara ideal (cámara de cajón) que mapea líneas rectas del espacio tridimensional en líneas rectas de la imagen.

Solamente estas informaciones son las condiciones necesarias para el enfoque monocular basado en líneas rectas, ver por ejemplo (Prescott, 1997). Sin embargo, la determinación de la función de distorsión es muy complicada en el caso de distorsiones radiales muy grandes.

El proceso de ajuste de las líneas rectas a las líneas muy curvadas es difícil y requiere tiempo largo. Por eso aplicamos otra idea.

El punto de inicio de nuestras investigaciones es una situación como en la imagen original de la figura 1.1 (arriba a la izquierda). Suponemos que no hay distorsiones tangenciales, como es una condición común (ver Prescott, 1997; Shih *et al.*, 1995; Shah, 1996; Weng *et al.*, 1992).

Entonces la primera meta es la determinación de la distorsión radial, es decir la determinación de la función de distorsión y del centro de distorsión.

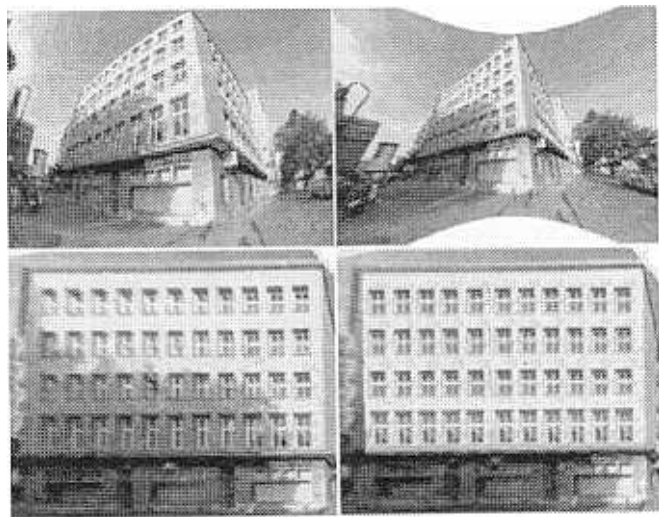


Fig. 1.1: Imagen original, rectificación no lineal, rectificación lineal, mejoramiento con ayuda de simetría y repetición

## 2 Eliminación de la Distorsión Radial

### 2.1 El Modelo de la Distorsión Radial

Para la descripción de la aplicación del mundo real a una imagen ideal usamos el modelo de la cámara de cajón. Además suponemos que nuestra imagen tiene pixeles

A partir de estos conocimientos es posible rectificar la imagen original de tal manera que ya no hay distorsiones radiales (ver la figura 1.1, arriba a la derecha, y la parte 2 del artículo).

Después tenemos que corregir las distorsiones proyectivas para obtener una aplicación homotética de una fachada seleccionada, con el resultado de una fachada correcta en tamaño y forma salvo un escalamiento (ver figura 1.1, abajo a la izquierda, y la parte 3 del artículo). Trabajos similares son descritos por Hemmleb (1997), Liebowitz (1998) y Van den Heuvel (1998b).

Al final mejoramos la imagen por uso de repetición de detalles (por ejemplo por uso de la repetición de las ventanas) y por consideración de la simetría de detalles (por ejemplo la simetría de las ventanas o la simetría de la fachada): En el ejemplo de la figura 1.1, se sustituyen los detalles mal aplicados (ventanas izquierdas) por los detalles bien aplicados (ventanas derechas), utilizando que estos detalles forman una retícula.

En el último cuadro de la figura 1.1 se muestra el resultado de nuestro método: La fachada izquierda del edificio está bien rectificada y los arquitectos pueden usar esta imagen para aplicaciones siguientes (vea la parte 4 del artículo).

isotrópicos, que la distorsión no lineal tiene solamente una parte radial y que el eje óptico es ortogonal al plano de la imagen (es decir, el centro de la distorsión coincide con el punto principal de la imagen).

Sean  $(x',y')$  las coordenadas de un punto  $p'$  de la imagen deformada,  $(x,y)$  las coordenadas del punto  $p$  correspondiente de la imagen ideal,  $P=P'=(X,Y)$  el punto principal de la imagen (ideal o deformada),  $r'$  la distancia entre  $p'$  y  $P$ ,  $r$  la distancia entre  $p$  y  $P$ .

Entonces la función de la distorsión  $r'=f(r)$  y su inversa  $r=f^{-1}(r')$  deben ser funciones monótonas con las propiedades

$$f(0)=f^{-1}(0)=0 \quad \lim_{r \rightarrow \infty} \frac{df(r)}{dr} = \lim_{r' \rightarrow \infty} \frac{df^{-1}(r')}{dr'} = 1$$

Puesto que las distorsiones  $|r'-f(r)|$  o  $|r-f^{-1}(r')|$  son pequeñas para distancias pequeñas  $r$  y  $r'$ , podemos representar las distorsiones por las ecuaciones siguientes con coeficientes  $a_i, b_i, c_i, d_i$  cualesquiera pequeños:

$$r' = r \left( 1 + a_1 r + a_2 r^2 + a_3 r^3 + \dots \right), \quad r = \frac{r'}{1 + b_1 r' + b_2 r'^2 + b_3 r'^3 + \dots}$$

Por otro lado, las expresiones para los radios  $r$  de la imagen ideal son

$$r = r' \cdot \left( 1 + c_1 r' + c_2 r'^2 + c_3 r'^3 \right), \quad r = \frac{r'}{1 + d_1 r' + d_2 r'^2 + d_3 r'^3}$$

Debido a que los coeficientes  $a_1, a_2, \dots$  son pequeñas, se pueden usar conversiones entre los coeficientes como en el ejemplo siguiente:

$$\begin{aligned} a_1 &= -b_1 \\ a_2 &= -b_2 + b_1^2 \\ a_3 &= b_1 b_2 - b_3 + (b_2 - b_1^2) b_1 \\ a_4 &= b_1 b_3 - b_4 + (b_2 - b_1^2) b_2 + (-2b_1 b_2 + b_3 + b_1^3) b_1 \end{aligned}$$

Por razón de la simetría radial de la distorsión, todos de los coeficientes del orden impar desaparecen ( $a_1=b_1=c_1=d_1=0$  etc.). En general, en los trabajos sobre la calibración se usan solamente los coeficientes del orden dos, es decir  $a_2, b_2, c_2, d_2$  (ver Devernay, 1995; Nomura, 1992; Shih *et al.*, 1995; Weng *et al.*, 1992) o los coeficientes del orden dos y cuatro, es decir  $(a_2, a_4), (b_2, b_4)$  etc. (ver Karras *et al.*, 1998, Prescott, 1997; Stein, 1999; Tsai, 1986).

Pero en el caso de las cámaras fish-eye no se puede describir la función de la distorsión como  $r'=r(1+a_2 r^2+a_4 r^4)$  ya

que tenemos distorsiones grandes. Por otro lado es muy complicado trabajar con más de dos coeficientes.

Por eso aquí nos basamos en un modelo con una función muy simple  $r=f^{-1}(r')$  que tiene una inversa exacta  $r'=f(r)$ . Aunque estas dos funciones son determinadas solamente por un parámetro  $c$ , las series de potencias tienen un número infinito de coeficientes:

$$\begin{aligned} r &= f^{-1}(r') = \frac{r'}{1 - c r'^2} \\ &= r' (1 + c r'^2 + c^2 r'^4 + c^3 r'^6 + \dots) \\ r' &= f(r) = \frac{\sqrt{1 + 4c r^2} - 1}{2c r} = \\ &= r (1 - c r^2 + 2c^2 r^4 - 5c^3 r^8 + \dots) \end{aligned}$$

En las figuras 2.1 y 2.2 se muestran las dos funciones  $r=f^{-1}(r')$  y  $r'=f(r)$  para el coeficiente  $c=5 \cdot 10^{-6}$ . Para valores  $r' > R=1/\sqrt{c}=447$  no hay valores  $r$ , es decir, todo el plano Euclideo (o la parte del espacio tridimensional con un ángulo del apertura de  $180^\circ$ ) está mapeado al interior de un "círculo límite" centrado en el punto principal y con el radio  $R=1/\sqrt{c}$  (ver figura 2.3). Por la rectificación de la distorsión radial este círculo límite se mapea en la línea infinita del plano proyectivo.

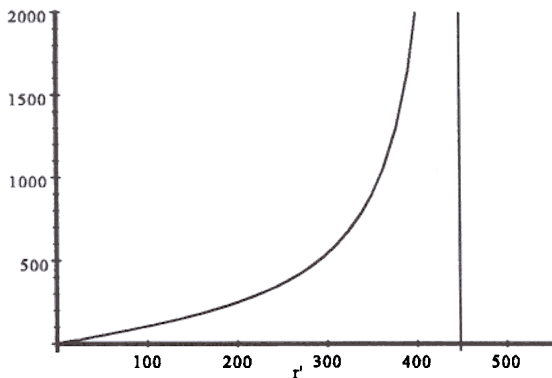


Figura 2. Función  $r=f^{-1}(r')$

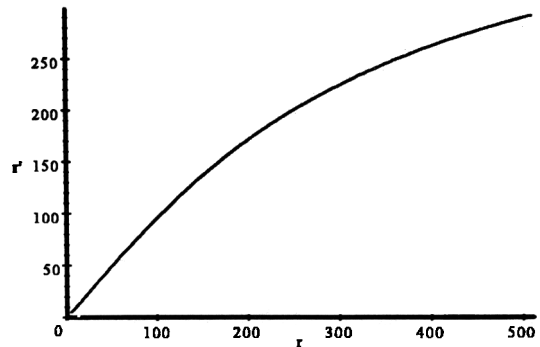


Figura 2.2: Función  $r'=f(r)$

Como ejemplo: el punto principal de la imagen como el centro de la distorsión y como el origen de los radios  $r$  y  $r'$  está localizado en la figura 3 cerca de la letra "L" de la palabra "CIRCULATION". Todos las líneas rectas de la imagen ideal que pasan por el punto principal, se preservan como líneas rectas en la imagen deformada. Puesto que  $(x-X)/(x'-X) = (y-Y)/(y'-Y) = r/r'$ , obtenemos de la formula (2.1) las coordenadas

$$\begin{aligned} x' - X &= (x - X) \cdot \frac{\sqrt{1 + 4c r^2} - 1}{2c r^2} \\ y' - Y &= (y - Y) \cdot \frac{\sqrt{1 + 4c r^2} - 1}{2c r^2} \end{aligned}$$

y

$$x-X = \frac{x'-X}{1 - \frac{r'^2}{R^2}}, \quad y-Y = \frac{y'-Y}{1 - \frac{r'^2}{R^2}} \quad (2.3)$$

con  $r^2=(x-X)^2+(y-Y)^2$  y  $r'^2=(x'-X)^2+(y'-Y)^2$ . En lo siguiente usamos siempre el radio  $R$  del círculo límite en lugar del parámetro  $c$  de la distorsión por que este radio  $R$  tiene inmediatamente un sentido geométrico.

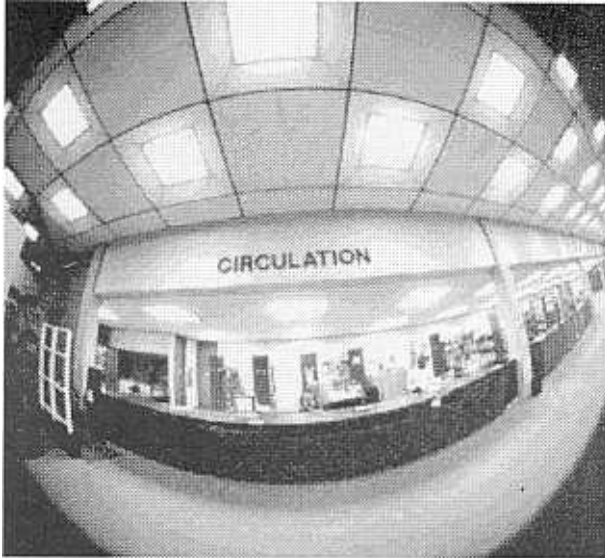


Figura 2.3: Ejemplo para un círculo límite

Una línea recta en la imagen ideal se describe por la ecuación  $(x-X)\cdot\cos\varphi+(y-Y)\cdot\sin\varphi=p$ , con  $p$  como distancia de la recta al punto principal. A partir de la fórmula (2.3) obtenemos

$$(x'-X)\cos\varphi + (y'-Y)\sin\varphi = p \cdot \left( \frac{r'^2}{R^2} \right) = \frac{p}{R^2} \cdot (R^2 - (x'-X)^2 - (y'-Y)^2)$$

que es la ecuación de un círculo en el plano  $x'-y'$ . De este modo, cada línea recta de la imagen ideal se mapea en un círculo de la imagen deformada (pero no al revés). Este círculo tiene el centro  $(\xi, \eta)$  con

$$\xi = X - \frac{\cos\varphi}{2cp} \quad \eta = Y - \frac{\sin\varphi}{2cp} \quad (2.5)$$

y el radio  $\rho$  con

$$\rho^2 = R^2 + (\xi - X)^2 + (\eta - Y)^2 \quad (2.6)$$

La fórmula (2.6) es nuestra "fórmula principal" de la rectificación de las distorsiones radiales.

## 2.2 La Detección de los Círculos en la Imagen

Nuestro objetivo ahora es detectar en la imagen radialmente distorsionada unos o todos los círculos que corresponden a líneas rectas en la imagen ideal. Para eso, en un primer paso buscamos segmentos rectos (mediante la determinación de contornos y su descomposición como secuencias de segmentos rectos). Después, a partir de cada pareja de segmentos rectos, determinamos el centro  $P_C$  de un círculo  $C$  como punto de intersección de las mediatrices, ver figura 2.4.

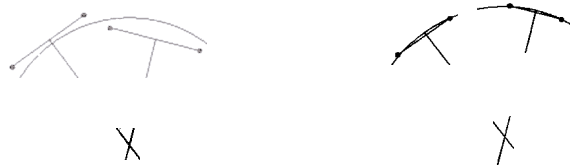


Figura. 2.4: Un círculo "no correcto" (izquierda) y un círculo "correcto" (derecha)

Si los puntos finales de los segmentos rectos resultan estar cerca de la periferia del círculo  $C$ , entonces llamamos este círculo "correcto" y lo usamos para los pasos siguientes. Reunimos todos los círculos correctos hallados en una lista  $L$ . Cabe notar que no todos los círculos de la lista  $L$  corresponden a líneas rectas en la imagen ideal (ver figura 2.5). En la siguiente sección nos dedicamos a la tarea de eliminar los círculos no correspondientes a líneas rectas originales. Al final quedan solamente círculos "adecuados" en la lista  $L$ , es decir, círculos que responden a líneas rectas en la imagen ideal.

## 2.3 El Proceso Iterativo y el Ajuste Homogéneo

En la sección anterior determinamos la lista  $L$  de todos los círculos correctos. Cada uno de estos círculos  $C_i$ ,  $i=1, \dots, N$  esta descrito por su centro  $(\xi_i, \eta_i)$  y su radio  $\rho_i$  y tiene que satisfacer la ecuación (2.6) donde  $(X, Y, R)$  son los parámetros del círculo límite. La idea es, determinar a partir de la lista  $L$  y la ecuación (2.6) y por un proceso iterativo un tupelo optimal  $(X, Y, R)$  para el círculo límite. En nuestro modelo, el centro  $(X, Y)$  del círculo límite debe que ser el punto principal de la imagen. Pero tanto  $(X, Y)$  como  $R$  son parámetros desconocidos.

En el primer paso del proceso iterativo, es decir en la fase de la "inicialización", estimamos como el punto principal el



punto medio  $(X_m, Y_m)$  de la imagen:  $(X^{[1]}, Y^{[1]}) = (X_m, Y_m)$ . Después calculamos a partir de ecuación (2.6) para cada uno

de los círculos  $C_i$  un radio  $R_i^{[1]}$  que satisface la ecuación (2.6). El valor medio de estos radios es

$$R^{[1]} = \frac{\sum_{i=1}^N p_i^{[1]} \cdot R_i^{[1]}}{\sum_{i=1}^N p_i^{[1]}} = \frac{\sum_{i=1}^N p_i^{[1]} \cdot \sqrt{\rho_i^2 - (\xi_i - X^{[1]})^2 - (\eta_i - Y^{[1]})^2}}{\sum_{i=1}^N p_i^{[1]}} \quad (2.7)$$

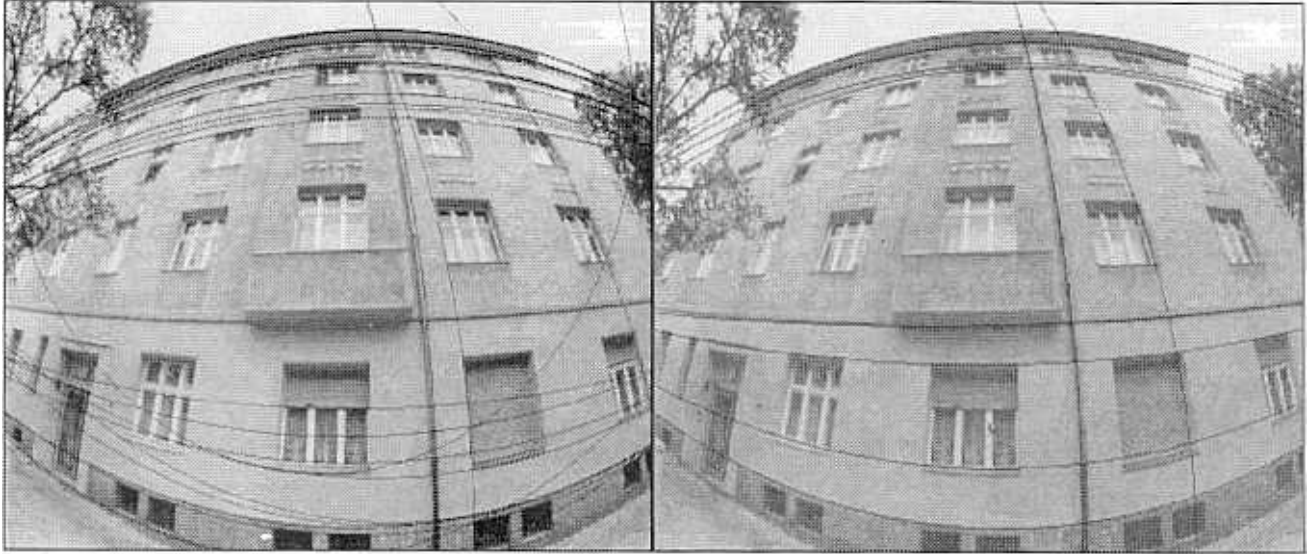


Figura 2.5: Todos los círculos correctos (izquierda) y los círculos adecuados (derecha)

Aquí usamos para cada uno de los círculos  $C_i$  un “peso”  $p_i$  que describe la seguridad con la cual el círculo  $C_i$  es procedente de una línea recta original. Inicialmente todos los pesos  $p_i^{[1]}=1$  son iguales. Si una expresión bajo la raíz es negativa, entonces el círculo  $C_i$  correspondiente recibe el peso  $p_i=0$ . La ecuación (2.7) es la primera parte del paso principal del proceso iterativo (la primera parte de la “actualización”).

En la segunda parte de la actualización calculamos para cada uno de los círculos  $C_i$  un radio  $\sigma_i$  que corresponde a los parámetros  $(X^{[1]}, Y^{[1]}, R^{[1]})$ :

$$\sigma_i = \sqrt{\rho_i^2 - (R^{[1]})^2} \quad (2.8)$$

Los círculos  $C_i^* = (\xi_i, \eta_i, \sigma_i)$  tienen que pasar teóricamente por el punto principal. En la práctica, sin embargo, estos círculos se intersectan solamente dentro de una región pequeña en la vecindad del punto principal (ver figura 2.6).

Ahora en la tercera parte de la actualización se determina un punto  $(X^{[2]}, Y^{[2]})$  como el punto óptimo de intersección de todos los círculos  $C_i^* = (\xi_i, \eta_i, \sigma_i)$ . Después la iteración empieza

nuevamente con la calculación de un radio  $R^{[2]}$  a partir de la ecuación (2.7). Los pesos  $p_i$  de los círculos se determinan en las siguientes iteraciones  $k=2,3,\dots$  por la ecuación

$$p_i^{[k]} = \left( \min \left( \frac{R_i^{[k-1]}}{R^{[k-1]}}, \frac{R^{[k-1]}}{R_i^{[k-1]}} \right) \right)^2$$

Aquí los parámetros  $R_i^{[k]}$  son los radios de los círculos obtenidos por la ecuación (2.7) en la  $k$ -ésima iteración.

De esta manera podemos determinar iterativamente a partir de la lista  $L$  de los círculos correctos el punto principal  $(X, Y)$  y el radio  $R$ :

$$\begin{aligned} & [(X_m, Y_m), L] \rightarrow [R^{[1]}, L] \rightarrow \\ & [(X^{[1]}, Y^{[1]}), L] \rightarrow [R^{[2]}, L] \rightarrow \dots \rightarrow [(X, Y, R)] \end{aligned}$$

Al final, los círculos  $C_i$  correctos de la lista  $L$  con un peso  $p_i$  suficientemente grande son círculos que corresponden a líneas

rectas en la imagen ideal. Estos círculos se llaman "círculos adecuados".

El número de los pasos de la iteración está en el rango 5-10. Nuestras investigaciones han mostrado que el error medio del radio  $R$  no es más de 4%. Para imágenes reales del tamaño  $900 \times 600$  píxeles (cámara Canon EF 15/2.8) hemos determinado un error medio maximal de las coordenadas del punto principal de  $\pm 16$  píxeles. También en los casos de imágenes parciales (entonces el punto principal no se encuentra en la proximidad del medio de imagen), podemos rectificar las

imágenes con una precisión suficiente para tareas en los campos de robótica o de la arquitectura. Los errores en la ventana central de la imagen ( $600 \times 400$  píxeles) no superan  $\pm 2$  píxeles.

El método aquí representado es numéricamente muy estable y también muy efectivo en respecto al tiempo. Por ejemplo: para 300 segmentos en la lista  $L$  extraído a una imagen de  $900 \times 600$  píxeles que resultan aproximadamente en 50 círculos adecuados necesita un PC con 1GHz solo un segundo. Este tiempo es 30-50 veces más rápido que el método descrito por Prescott (1997).

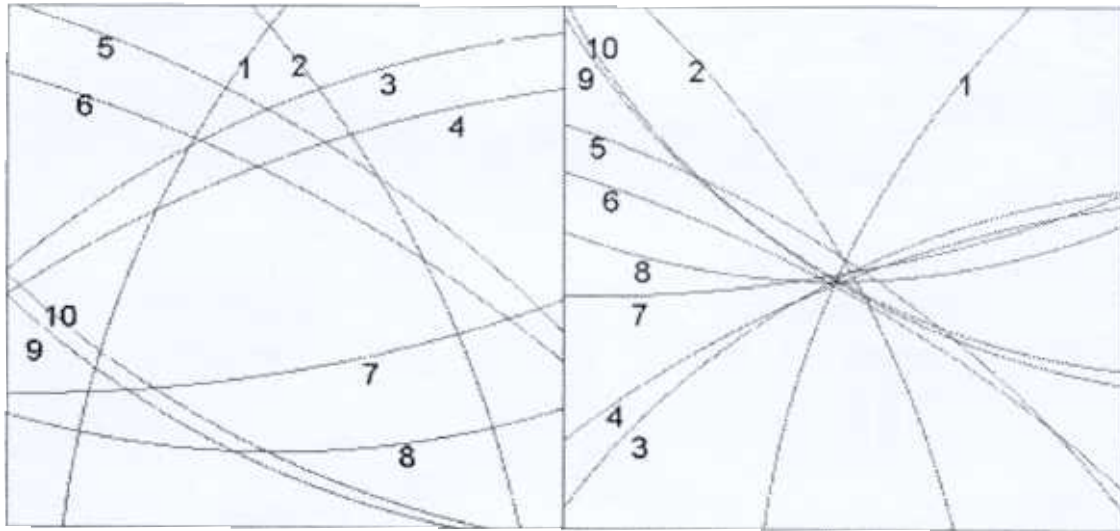


Figura 2.6: Los círculos adecuados  $C_i$  con radios  $\rho_i$  (izquierda) y los círculos  $C_i^*$  con radios  $\sigma_i$ , que tienen que intersectarse teóricamente en el punto principal (derecha)

### 3 Rectificación Lineal

#### 3.1 El Modelo de la Cámara Ideal

Las ecuaciones del mapeo central-proyectiva son

$$\xi = \frac{a_{10} + a_{11}x + a_{12}y + a_{13}z}{a_{30} + a_{31}x + a_{32}y + a_{33}z} = \frac{a_{10} + \langle \mathbf{a}_1, \mathbf{r} \rangle}{a_{30} + \langle \mathbf{a}_3, \mathbf{r} \rangle}$$

$$\eta = \frac{a_{20} + a_{21}x + a_{22}y + a_{23}z}{a_{30} + a_{31}x + a_{32}y + a_{33}z} = \frac{a_{20} + \langle \mathbf{a}_2, \mathbf{r} \rangle}{a_{30} + \langle \mathbf{a}_3, \mathbf{r} \rangle}$$

Los parámetros  $a_{ik}$  describen la proyección. El vector tridimensional  $\mathbf{r}$  y su vector transpuesto  $\mathbf{r}^T = (x, y, z)$  representan un punto espacial cuyo imagen bajo la proyección central es el punto de imagen  $(\xi, \eta)$ . Aquí, las expresiones  $\langle \mathbf{a}_i, \mathbf{r} \rangle$  denotan los productos escalares entre los dos vectores  $\mathbf{a}_i$  y  $\mathbf{r}$ .

Todos los puntos espaciales de la línea recta  $\mathbf{r} = \mathbf{c} + \lambda \mathbf{t}$ , donde  $\mathbf{c}$  es el centro de proyección y  $\mathbf{t}$  es el vector direccional del "rayo de vista", son puntos originales para el punto de imagen  $(\xi, \eta)$ . Las siguientes fórmulas valen:

$$\mathbf{c} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} a_{10} \\ a_{20} \\ a_{30} \end{pmatrix} - \begin{pmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \mathbf{a}_3^T \end{pmatrix} \mathbf{a}_0$$

para la calculación del centro de proyección  $\mathbf{c}$  y

$$\mathbf{t}_{\xi, \eta} = \mathbf{a}_1 \times \mathbf{a}_2 + \xi \cdot \mathbf{a}_2 \times \mathbf{a}_3 + \eta \cdot \mathbf{a}_3 \times \mathbf{a}_1 \quad (3.1)$$

para la determinación del vector  $\mathbf{t}_{\xi, \eta}$  a partir de los vectores  $\mathbf{a}_i$  y el punto de imagen  $(\xi, \eta)$ .

Los cuatro vectores  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$  y  $\mathbf{a}_0^T = (a_{10}, a_{20}, a_{30})$  son llamados los "vectores de la cámara" (ver por ejemplo WWW\_1).

En la figura 3.1 se muestra el esquema de la proyección central. Para una cámara ideal el punto principal  $P = (X, Y)$  es la proyección ortogonal del centro de la proyección al plano de imagen. Sea  $D$  una línea recta espacial como el conjunto de todos los puntos espaciales  $\mathbf{r} = \mathbf{r}_D + \lambda \mathbf{t}$  con  $\mathbf{t}$  como vector direccional de esta línea.

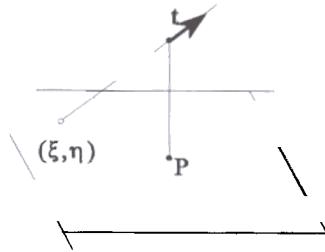


Figura 3.1: Punto de imagen  $(\xi, \eta)$  y rayo de vista

Entonces todos los puntos de imagen  $(\xi_\lambda, \eta_\lambda)$  correspondientes a los puntos de la línea  $D$ , con  $-\infty < \lambda < +\infty$ , son dados por las fórmulas

$$\xi_\lambda = \frac{a_{10} + \langle \mathbf{a}_1, \mathbf{r}_D + \lambda \mathbf{t} \rangle}{a_{30} + \langle \mathbf{a}_3, \mathbf{r}_D + \lambda \mathbf{t} \rangle}, \quad \eta_\lambda = \frac{a_{20} + \langle \mathbf{a}_2, \mathbf{r}_D + \lambda \mathbf{t} \rangle}{a_{30} + \langle \mathbf{a}_3, \mathbf{r}_D + \lambda \mathbf{t} \rangle}$$

Para el límite  $\lambda \rightarrow \infty$  obtenemos el punto de imagen  $F_t = (\xi_\infty, \eta_\infty)$  con

$$\xi_\infty = \lim_{\lambda \rightarrow \infty} \frac{a_{10} + \langle \mathbf{a}_1, \mathbf{r}_D + \lambda \mathbf{t} \rangle}{a_{30} + \langle \mathbf{a}_3, \mathbf{r}_D + \lambda \mathbf{t} \rangle} = \frac{\langle \mathbf{a}_1, \mathbf{t} \rangle}{\langle \mathbf{a}_3, \mathbf{t} \rangle}$$

$$\eta_\infty = \frac{\langle \mathbf{a}_2, \mathbf{t} \rangle}{\langle \mathbf{a}_3, \mathbf{t} \rangle}$$

como "punta de fuga" (*vanishing point*) de la línea  $D$ . Todas las líneas paralelas a  $D$  tienen el mismo punto de fuga ya que  $(\xi_\infty, \eta_\infty)$  no depende del punto espacial  $\mathbf{r}_D$ . En otras palabras, cada vector direccional  $\mathbf{t}$  corresponde únicamente a un punto de imagen y viceversa (si  $\mathbf{t}$  está en un plano paralelo al plano de imagen, entonces  $(\xi_\infty, \eta_\infty)$  es un punto en el infinito).

Cada "base ortogonal"  $B_{\text{orto}}$ , descrito por tres direcciones ortogonales de líneas rectas, define tres puntos de fuga. Estas situaciones hay por ejemplo en el caso de la arquitectura (ver figuras 3.2, 3.3 y 3.4).

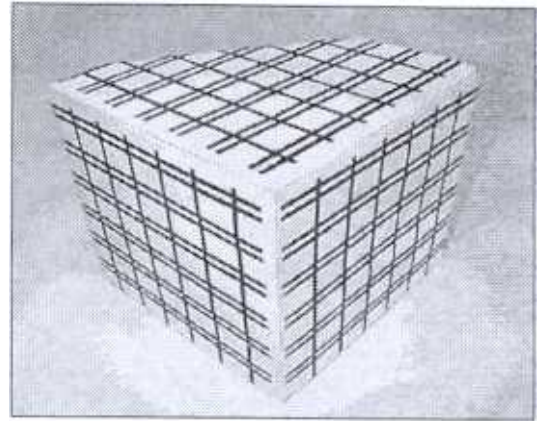


Figura 3.2: Un "cuerpo arquitectónico" con tres lados ortogonales

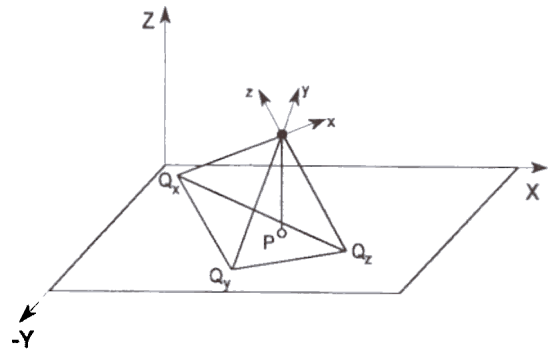


Figura 3.3: Tres puntos de fuga  $Q_x, Q_y, Q_z$

El punto principal  $P$  conjuntamente con los puntos de fuga  $Q_x, Q_y, Q_z$  forma un sistema ortocentral ("orthocentric system", ver Weinstein (1999)), donde las tres alturas del triángulo  $T = (Q_x, Q_y, Q_z)$  se intersectan en  $P$ . Este sistema ortocentral tiene las propiedades siguientes:

- El punto principal  $P$  está localizado dentro de  $T$ .
- La distancia focal  $f$  (la distancia entre el centro de proyección  $c$  y el plano de imagen) determina el tamaño del triángulo  $T$ .
- Una rotación de  $B_{\text{orto}}$  alrededor del eje óptico (la línea determinada por  $c$  y  $P$ ) proporciona solamente una rotación de  $T$ .
- Una traslación de  $P$  sólo desplaza el triángulo  $T$ .

Los vectores  $\mathbf{t}_x, \mathbf{t}_y, \mathbf{t}_z$  de los rayos de vista determinados por los puntos de fuga  $Q_x = (\zeta_x, \eta_x)$ ,  $Q_y = (\zeta_y, \eta_y)$ ,  $Q_z = (\zeta_z, \eta_z)$  y el centro de proyección  $c$  son dados por las fórmulas

$$\begin{pmatrix} \xi_x - X \\ \eta_x - Y \\ -f \end{pmatrix} = \mathbf{t}_x = \begin{pmatrix} \xi_y - X \\ \eta_y - Y \\ -f \end{pmatrix} = \mathbf{t}_y = \begin{pmatrix} \xi_z - X \\ \eta_z - Y \\ -f \end{pmatrix} = \mathbf{t}_z$$

A partir de las tres condiciones de ortogonalidad

$$\langle \mathbf{t}_x, \mathbf{t}_y \rangle = 0 \quad , \quad \langle \mathbf{t}_y, \mathbf{t}_z \rangle = 0 \quad , \quad \langle \mathbf{t}_z, \mathbf{t}_x \rangle = 0$$

podemos determinar las tres incógnitas, es decir el punto principal  $P=(X,Y)$  y la distancia focal  $f$ .

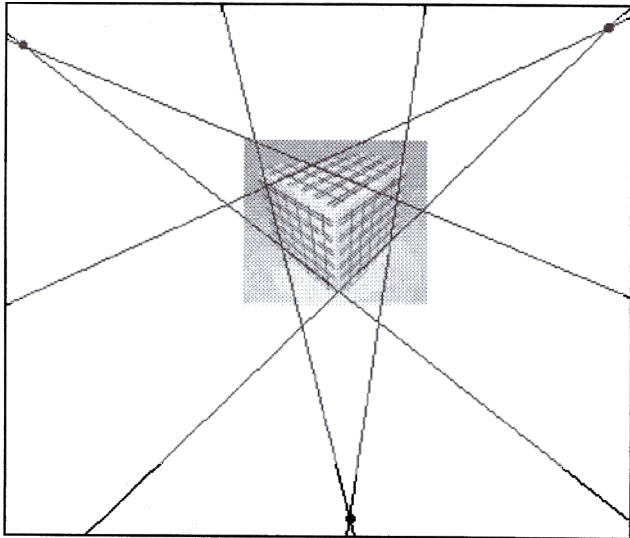


Figura 3.4: Un objeto "arquitectónico" con sus tres puntos de fuga para las tres direcciones ortogonales

### 3.2 El Modelo de la Rectificación Lineal

Cada rotación tridimensional de la cámara en torno su centro de proyección  $c$  descrito por una matriz de rotación  $\mathbf{R}$  proporciona una transformación  $\mathbf{T}_R$  bidimensional de la imagen original. El razón es que todas las líneas rectas del espacio que pasan por el centro de la proyección son mapeadas únicamente en puntos de la imagen.

Para cada uno de los puntos de imagen  $(\xi, \eta)$  existe un y sólo un rayo de vista  $\mathbf{t}_{\xi, \eta}$  (ver ecuación (3.1)). Hay una rotación especial  $\mathbf{R}_k(\xi, \eta)$  dada por la rotación del rayo de vista  $\mathbf{t}_{\xi, \eta}$  al eje óptico. Esta rotación determina una transformación proyectiva  $\mathbf{T}_k(\xi, \eta)$  de la imagen, la cual es llamada "transformación de Kanatani" y fue descrita por primera vez por Kanatani (1988), ver también Voss, (1995a) y WWW\_1.

Una propiedad importante de esta transformación es que cada plano que es inicialmente ortogonal al rayo de vista  $\mathbf{t}_{\xi, \eta}$  es paralelo al plano de imagen después de la transformación  $\mathbf{T}_k(\xi, \eta)$ . El mapeo de un plano espacial  $P_{\text{espacio}}$  paralelo al plano de imagen  $P_{\text{imagen}}$  es una transformación homotética de  $P_{\text{espacio}}$  a  $P_{\text{imagen}}$  - por eso cada distorsión proyectiva es eliminada.

Las fórmulas para la transformación de Kanatani son las siguientes. Sea  $Q=(\xi_Q, \eta_Q)$  el punto de fuga que se debe

transformar en el punto principal  $P=(X,Y)$ . En el primero paso giramos la imagen de tal manera que cada punto  $(\xi, \eta)$  se transforma en un punto  $(\xi^*, \eta^*)$ :

$$\alpha = \arcsin \frac{\eta_Q - Y}{\sqrt{(\xi_Q - X)^2 + (\eta_Q - Y)^2}}$$

$$\xi^* = (\xi - X) \cos \alpha + (\eta - Y) \sin \alpha + X$$

$$\eta^* = -(\xi - X) \sin \alpha + (\eta - Y) \cos \alpha + Y$$

Entonces calculamos las nuevas coordenadas

$$\xi' = X - f \cdot \tan \left( \frac{\arctan |\xi_Q - X|}{f} + \frac{\arctan |\xi^* - X|}{f} \right)$$

$$\eta' = Y + (\eta^* - Y) \frac{\sqrt{f^2 + (\xi' - X)^2}}{\sqrt{f^2 + (\xi^* - X)^2}}$$

De esta manera podemos transformar por ejemplo la imagen de una fachada la cual ya no tiene distorsiones no lineales pero todavía tiene distorsiones proyectivas a una imagen "rectificada" en la cual la fachada es mostrada de frente (ver figuras 3.5 y 1.1).

En la figura 3.4 se muestran los tres puntos de fuga. El punto principal se localiza cerca del centro de la imagen. Si se determina por ejemplo el vector direccional  $\mathbf{t}_x$  para el punto de fuga  $Q_x$  y se gira el rayo de vista correspondiente por la transformación de Kanatani en el eje óptico, entonces cada plano  $P_{\text{objeto}}$  del objeto que es ortogonal al vector  $\mathbf{t}_y \times \mathbf{t}_z$  se transforma en un plano paralelo al plano de imagen. Por esta paralelidad el plano  $P_{\text{objeto}}$  aparece en la imagen transformada como una reproducción a escala (ver figura 3.5). La precisión de la rectificación (dependiente de la calidad de las imágenes originales) nos proporciona un cociente ente las longitudes de los dos lados de un cuadrado de 0.97 - 1.03 (en lugar de 1.00 en caso de una reconstrucción ideal).

### 3.3 La Determinación de los Puntos de Fuga

Para la rectificación lineal de sección 3.2 la determinación precisa de los puntos de fuga es de gran importancia. Por eso hay muchos trabajos sobre esta tarea en la literatura (Caprile, 1990; Gamba *et al.*, 1996; Lutton *et al.*, 1994; McLean, 1995; Palmer, 1993; Van den Heuvel, 1998a; Shufelt, 1999).

En muchos casos se usan métodos basados en la representación de la esfera de Gauss (*Gaussian sphere method* GSM, ver por ejemplo Lutton (1994). Pero este método tiene unas desventajas por la cuantización (discretización) necesaria de la esfera a partir del método de Hough.



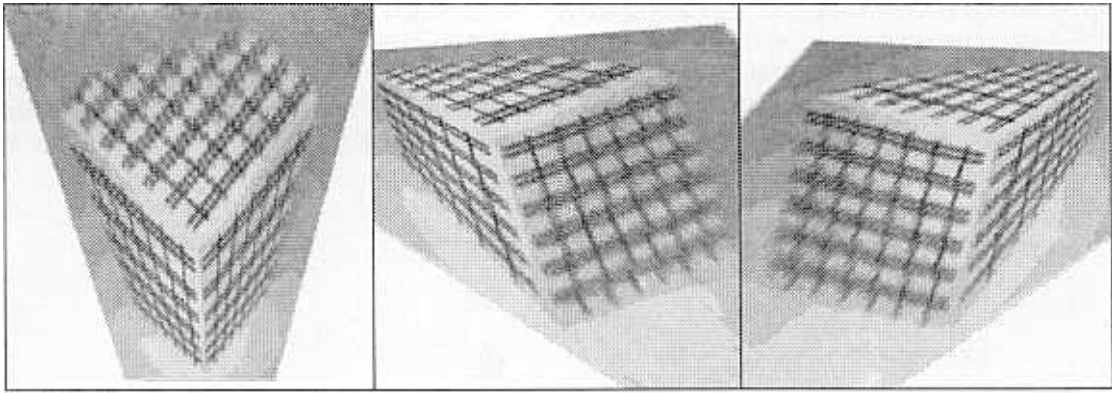


Figura 3.5: Los tres lados de un objeto "arquitectónico" como las reproducciones a escala obtenidas por la transformación de Kanatani

Por eso queremos presentar aquí dos nuevos métodos geométricos para la determinación robusta de los puntos de fuga (ver Bräuer-Burchardt (2000)). Estos métodos son el enfoque a partir del círculo de Thales (TCM) y el enfoque con

punto  $M_c$  está localizado en media de la línea entre A y el punto de fuga Q con  $A M_c = M_c Q = R_c$  así que entonces también conocemos Q. Este enfoque TCM es más simple y más robusto que el método de la esfera de Gauss (ver figura 3.7).

Aún mejor es el enfoque de la minimización de las áreas de triángulos. De nuevo partimos de una lista de segmentos rectos. Si un segmento  $P_1P_2$  pasa a través del punto de fuga Q, entonces el triángulo  $P_1P_2Q$  tiene un área igual a zero:

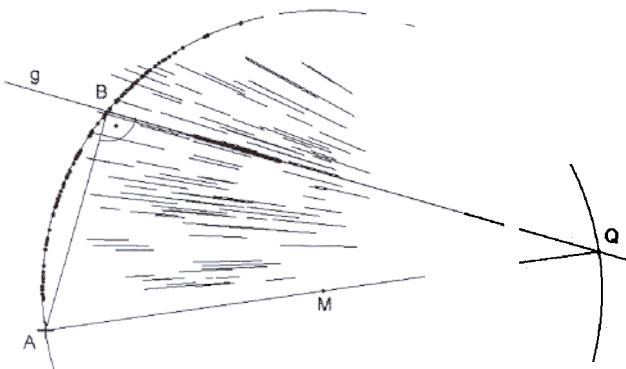


Figura 3.6: Segmentos, punto de base B y círculo de Thales que pasa por todos los puntos de base

ayuda de la minimización de las áreas de triángulos (TAM). Para ambos enfoques usamos una lista de segmentos rectos  $S = ((x_1, y_1), (x_2, y_2))$  los cuales son extraídos de una imagen sin distorsiones no lineales.

Cada segmento S determina una línea recta g. Ahora buscamos el punto de base B en esta línea con respecto a un punto arbitrario A (ver figura 3.6, por razones de experiencias prácticas se recomienda escoger el punto A verca de la nube de segmentos).

Todos los puntos de base B sobre las líneas representadas por los segmentos de la lista con respecto al punto A están en el caso ideal sobre un círculo, el círculo de Thales, ya que el ángulo entre la altura AB y la línea BQ siempre es un ángulo recto. Si determinamos un círculo óptimo C a través de los puntos B, entonces conocemos su centro  $M_c$  y su radio  $R_c$ . El

$$\text{área}(P_1P_2Q) = \frac{1}{2} \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_Q & y_Q & 1 \end{vmatrix} = 0$$

Podemos escribir este área también como

$$x_1y_2 - x_1y_Q - x_2y_1 + x_2y_Q + x_Qy_1 - x_Qy_2 = 0$$

Pero a partir de los errores de medición estas áreas no son zero, sino tenemos valores pequeños  $\delta_i$  para cada segmento  $S_i$ . Por eso la solución de la tarea

$$F(x_Q, y_Q) = \sum_{i=1}^n \left[ x_1^{[i]} y_2^{[i]} - x_2^{[i]} y_1^{[i]} - (x_1^{[i]} - x_2^{[i]}) y_Q + (y_1^{[i]} - y_2^{[i]}) x_Q \right]^2 \rightarrow \text{mínimo}$$

nos proporciona a partir de  $\partial F / \partial x = \partial F / \partial y = 0$  dos ecuaciones lineales para las dos incógnitas  $x_Q$  y  $y_Q$ . Si los puntos terminales de los segmentos tienen errores con una desviación estándar de 0.1 - 1.0 píxeles, entonces podemos determinar la distancia del punto de fuga al punto principal con una precisión con un error de 0.5 - 5.0 % y la dirección al punto

de fuga tiene solamente errores de 0.02 - 0.25 grados donde el enfoque TAM es un poco mejor que el enfoque TCM (ver figura 3.7 y Bräuer.Burchardt (2000)). En la parte izquierda de la figura 3.7 se muestra el error porcentual de la distancia entre el punto principal de la imagen y el punto de fuga para

unas cientos de simulaciones como la función de la desviación estándar de las coordenadas de los puntos finales de los segmentos (de 0.1 hasta 1.0 píxeles). En la parte derecha se muestra el error de la dirección del punto principal al punto de fuga en grados (también como función de las coordenadas).

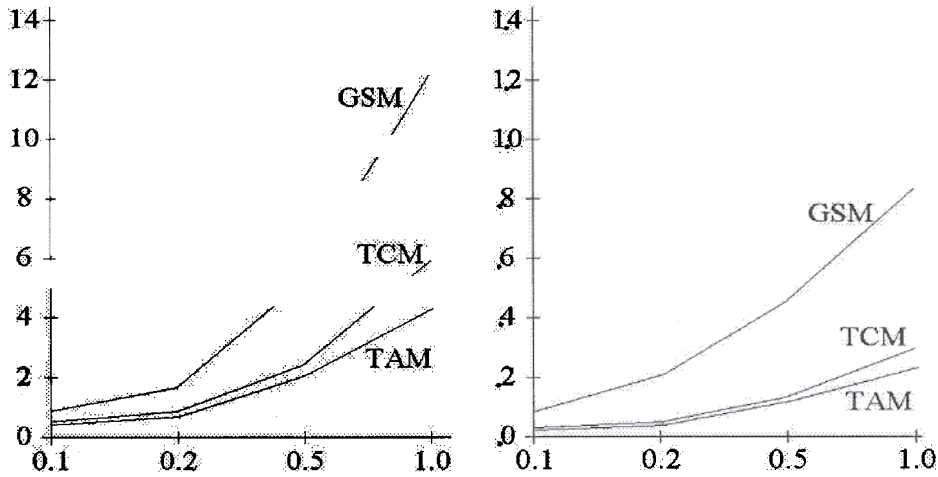


Figura 3.7: Errores de los enfoques TCM y TAM en comparación con el método de la esfera de Gauss GSM

## 4 Mejoramiento de las Imágenes

### 4.1 El Método SDR

Hemos desarrollado un método nuevo (*shift detection by restoration*, SDR) tanto para el apareamiento de imágenes como para la detección de desplazamientos entre dos imágenes (Voss, 1999a; Voss, 1999b; Suesse, 1999). La idea principal es la siguiente: Sean dadas dos imágenes I e I' del mismo tamaño y sea D<sub>0</sub> el impulso unitario (función de Dirac) que tiene en todas partes de la imagen los valores de gris 0 y solamente en el origen el valor 1. Ahora, si S<sub>u,v</sub> es el operador de desplazamiento que debe realizar un desplazamiento entero u en dirección X y un desplazamiento entero v en dirección Y, entonces obtenemos para el desplazamiento cíclico la fórmula

$$I' = S_{u,v}I = S_{u,v}(I * D_0) = I * (S_{u,v}D_0) = I * D_{u,v} \quad (4.1)$$

donde el símbolo \* representa la convolución entre funciones y D<sub>u,v</sub> es una imagen de impulso unitario desplazado.

Si aplicamos a esta ecuación (4.1) la transformada de Fourier  $\mathcal{F}$ , entonces obtenemos por el teorema de convolución la ecuación  $\mathcal{F}(I') = \mathcal{F}(I * D_{u,v}) = \mathcal{F}(I) \cdot \mathcal{F}(D_{u,v})$ . Por eso tenemos para los coeficientes de Fourier  $\mathcal{F}_{k,l}(I')$  de la función I' las

ecuaciones siguientes:

$$\mathcal{F}_{k,l}(I') = \mathcal{F}_{k,l}(I) \cdot \mathcal{F}_{k,l}(D_{u,v}^*)$$

Formalmente obtenemos los coeficientes de la transformada de Fourier  $\mathcal{F}_{k,l}(D_{u,v})$  de la imagen del impulso unitario desplazado D<sub>u,v</sub> por  $\mathcal{F}_{k,l}(I') / \mathcal{F}_{k,l}(I)$ . Pero esta solución no es calculable si unos coeficientes  $\mathcal{F}_{k,l}(I)$  de la función transformada  $\mathcal{F}(I)$  son igual de cero. También en el caso de que todos los coeficientes  $\mathcal{F}_{k,l}(I)$  son diferente de cero hay dificultades. Si unos coeficientes  $\mathcal{F}_{k,l}(I)$  son pequeños, entonces errores pequeños en los valores de gris de la imagen I (causados por ruido) resultan en errores grandes para los coeficientes D<sub>u,v</sub>.

Para eliminar estas dificultades usamos la teoría de Wiener (ver Pratt (1991) y Gonzales (1993)) que proporciona la fórmula de la restauración

$$\mathcal{F}_{k,l}(D_{u,v}) = \frac{\mathcal{F}_{k,l}^*(I) \cdot \mathcal{F}_{k,l}(I')}{\mathcal{F}_{k,l}^*(I) \cdot \mathcal{F}_{k,l}(I) + \gamma} \quad (4.2)$$

para los coeficientes de Fourier  $\mathcal{F}_{k,l}(D_{u,v})$  de la transformada del impulso desplazado restaurado. El símbolo  $\mathcal{F}_{k,l}^*$  describe el complejo conjugado de  $\mathcal{F}_{k,l}$  y  $\gamma$  es una constante (dependiente del ruido y de las perturbaciones). En otras palabras: la transformada inversa  $\mathcal{F}^{-1}(\mathcal{F}(D))$  es la solución restaurada de la ecuación (4.1), la así llamada "pseudo inversa".



Figura 4.1: Detección de un pico a partir de fórmula (4.2)

De fórmula (4.2) podemos obtener las soluciones comunes para la detección de un desplazamiento. Con  $\gamma \rightarrow 0$  obtenemos la solución directa  $\mathfrak{F}(D) = \mathfrak{F}(I') / \mathfrak{F}(I)$  y para  $\gamma \rightarrow \infty$  obtenemos (salvo una constante) la correlación cruzada

$$\mathfrak{F}_{k,l}(D_{u,v}) = \mathfrak{F}_{k,l}^*(I) \cdot \mathfrak{F}_{k,l}(I'),$$

es decir  $D_{u,v} = I_{\text{refl}} * I'$  donde  $I_{\text{refl}}$  denota la imagen reflejada (reflexión por el origen).

En la figura 4.1 se muestra la potencia del método SDR. Las tres imágenes de la columna izquierda son desplazada cíclicamente (en la columna media arriba), desplazada con perturbaciones en los bordes (en medio) y desplazada con perturbaciones y ruido (columna media abajo).

Las imágenes de los picos detectados (indicado por flechas) usando la fórmula (4.2) son muestras en la columna derecha. Los picos que describen los desplazamientos son muy destacados y se puede detectar fácilmente estos picos.

## 4.2 La Detección de las Simetrías

Aplicamos el método SDR para la detección de simetrías en imágenes. En figura 4.2 mostramos una imagen original (arriba a la izquierda, Palacio Nacional de México). Si se refleja esta imagen  $I_0$  por el eje vertical central de la imagen, se recibe la imagen reflejada  $I_r$  (arriba en medio). A partir de la simetría de la imagen original tenemos un desplazamiento entre ambas imágenes  $I_0$  e  $I_r$ .

Este desplazamiento  $d_2$  es el doble de la distancia  $d_1$  entre el eje vertical central y el eje de la simetría en los imágenes. Por eso se obtiene una imagen  $D_{u,v}$  del impulso unitario desplazado (arriba a la derecha). Este impulso tiene una distancia  $d_2$  del eje central de la imagen.

El lugar del pico del impulso unitario determina el eje de simetría. Ahora podemos determinar una línea de simetría en la imagen original (abajo a la izquierda) y en la imagen reflejada (abajo en medio). Si se desplaza las imágenes  $I_0 \rightarrow I'_0$  y  $I_r \rightarrow I'_r$  de tal manera que los ejes de simetría son iguales (figura 4.2 abajo a la izquierda y en medio), entonces grandes partes de las imágenes son iguales. En nuestro ejemplo se detectó bien la simetría presente como muestran las pequeñas diferencias  $|I'_0 - I'_r|$  de ambas imágenes desplazadas (abajo a la derecha).

Para tareas arquitectónicas podemos aplicar este método de la detección de la simetría para la reconstrucción de las partes de un edificio que no son visibles en la imagen original (por supuesto si sabemos que estas partes son partes reflejadas).

## 4.3 La detección de repeticiones

Si tenemos una imagen  $I$  con un objeto pequeño, entonces podemos generar una imagen de repetición  $I'$  de este objeto por la fórmula siguiente:



$$I' = \left( \sum_r \sum_s D_{r,a,s,b} \right) I = D^* I$$

Aquí  $D_{r,a,s,b}$  es el impulso unitario desplazado por  $r \cdot a$  píxeles en dirección X y por  $s \cdot b$  píxeles en dirección Y, donde  $a$  es la distancia de los impulsos en dirección X y  $b$  la distancia de los impulsos en dirección Y. De esta manera construimos una imagen  $I'$  con ayuda del operador  $D^*$  en la cual tenemos muchas repeticiones del objeto pequeño original.

A partir de la fórmula (4.2) podemos determinar aproximadamente el operador  $D^*$  como suma de impulsos unitarios separados por la restauración de Wiener, es decir, podemos determinar la rejilla de las repeticiones del objeto pequeño. En figura 4.3 se muestra un ejemplo. Una ventana bien mapeada es nuestro "objeto pequeño" (rectángulo negro pequeño en la imagen derecha).

La restauración nos da la rejilla de los impulsos (en medio), y entonces podemos sustituir las ventanas mal representadas por el objeto pequeño (la ventana "buena"). De esta manera obtenemos el resultado de las figuras 4.3 y 1.1.

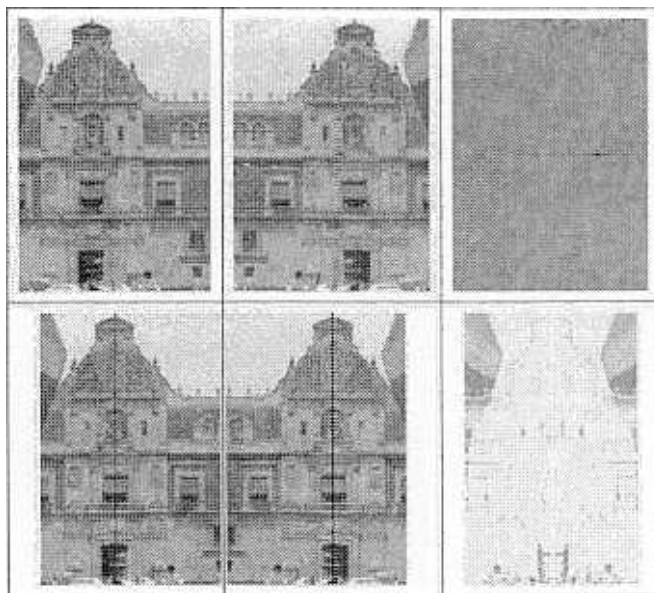


Figura 4.2: Arriba - imágenes  $I_0$  y  $I_r$ , imagen del impulso  $D_{u,v}$ , abajo - imágenes desplazadas y su diferencia  $|I'_0 - I_r|$

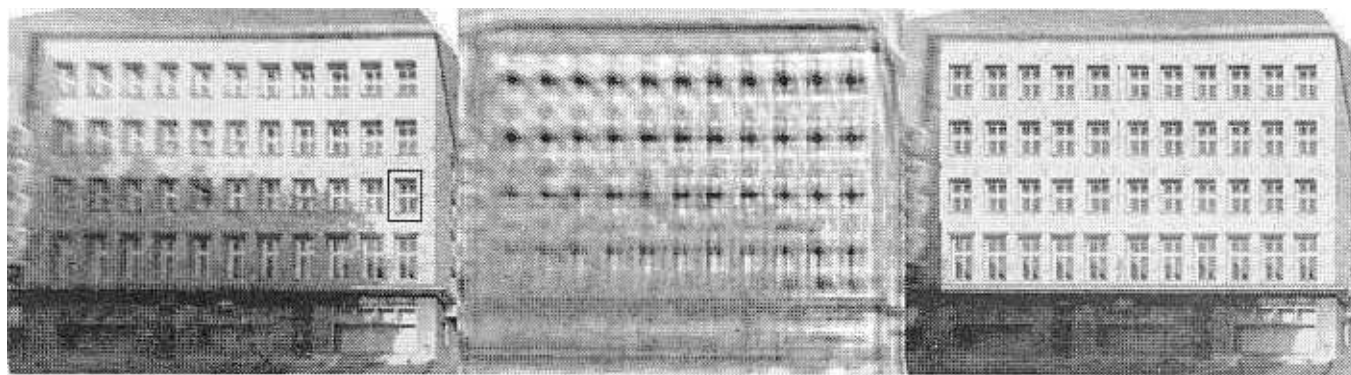


Figura 4.3: Fachada original y objeto pequeño (izquierda), imagen de los impulsos  $D^*$  (en medio) y fachada mejorada que es determinada por repetición (derecha)

## 5 Conclusiones

Hemos mostrado como podemos remover las distorsiones radiales en imágenes captadas por cámaras de tipo fish-eye aplicando un método nuevo y simple (parte 2 de este trabajo). Las ventajas de este método son que se puede aplicar también en casos de imágenes monoculares y si no tenemos datos sobre la cámara usada.

Solo con la información a priori de que hay líneas rectas en los objetos mapeados, podemos determinar el punto principal de la imagen (al mismo tiempo el centro de la distorsión radial) y el parámetro de la distorsión (que sustitue

en nuestro método dos o más parámetros necesarios en otros enfoques).

Este nuevo método puede usarse automáticamente, ya que los pasos necesarios (detección de segmentos lineales, determinación de círculos, evaluación de los círculos detectados, determinación del punto principal y del parámetro de la distorsión, rectificación de la imagen) son muy estables contra ruido, sombra, translaciones, rotaciones, escalaciones y aún transformaciones afines y proyectivas (en todos estos casos líneas rectas se conservan).

Mediante la eliminación de la distorsión radial (si es necesaria) obtenemos imágenes sin distorsiones no lineales



pero todavía con distorsiones lineales (distorsiones proyectivas). Podemos utilizar nuevamente la información a priori sobre los objetos mapeados por ejemplo sobre la paralelidad de líneas rectas en la realidad. Estas propiedades comunmente estan satisfechas para objetos tecnicos y para objetos arquitectónicos. En la imagen, las líneas paralelas del espacio aparecen como líneas rectas que se cortan aproximadamente en el punto de fuga. Con ayuda de tres puntos de fuga para tres tuplos ortogonales de líneas rectas podemos determinar nuevamente el punto principal  $(X, Y)$  de nuestra imagen y la distancia focal  $f$ .

Finalmente podemos transformar las partes de la imagen que representan planos espaciales (por ejemplo fachadas de edificios) de tal manera que estas partes resultan ser fieles copias de los planos espaciales salvo una escala. Ahora todas las distorsiones lineales (proyectivas y afines) son eliminadas así que mediciones métricas o comparaciones con dibujos técnicos son posibles (parte 3). Esta eliminación de las distorsiones lineales se puede ejecutar automaticamente así que el método es aplicable para robots móviles.

En caso de imágenes arquitectónicas podemos mejorar los resultados de la rectificación lineal a partir del método SDR. La información a priori sobre simetrías de los objetos mapeados y sobre repeticiones de detalles pequeños en un objeto (por ejemplo ventanas de una fachada) permite un mejoramiento de las imagenes (parte 4). En muchos casos este mejoramiento se puede realizar casi automaticamente con solo pocas intervenciones manuales.

## Referencias

**Albertz, J., A. Wiedemann.** "Acquisition of CAD data from existing buildings by photogrammetry". In: P.J.Pahl, H.Werner (eds.), Balkema, Rotterdam & Brookfield 1995, pp. 859-866

**Bookstein, F.L.** "Fitting conic sections to scattered data". *Computer Graphics and Image Processing* 9 (1979) 56-71.

**Brown, D.C.** "Close-range camera calibration" *Photogram. Eng.* 37 (1971) 855-866

**Bräuer-Burchardt, C., K. Voss.** "Robust vanishing point determination in noisy images". *Proc. 15th IAPR Conf., Barcelona Sept.2000, Vol.I, pp.559-562*

**Caprile, B., V.Torre.** "Using vanishing points for camera calibration". *Proc. IJCV 1990, pp. 127-140*

**Devernay, F., O.Faugeras.** "Automatic calibration and removal of distortion from scenes of structured environments". *SPIE* 2567 (1995) 62-72

Con métodos fotogramétricos se puede realizar precisiones de medición determinadas por el equipamiento técnico (cámaras, cuerpos de calibración etc.). En contra nuestro enfoque se basa en informaciones que se pueden extraer de las imágenes investigadas o en informaciones a priori sobre las propiedades de los objetos mapeados (linealidad, ortogonalidad etc.). No necesitamos informaciones sobre los parámetros interiores de la cámara (distancia focal, posición del punto principal etc.), parámetros exteriores de la cámara (posición y orientación de la cámara), puntos de referencia de los objetos tridimensionales y no usamos cuerpos de calibración.

Por eso, nuestros métodos no son equivalentes con métodos fotogramétricos estandares. Pero los beneficios de nuestros métodos consisten en la aplicabilidad también en casos si hay solamente imágenes únicas. Por ejemplo es posible reconstruir edificios históricos destruidos o medir objetos arquitectónicos bajo la suposición de uso de métodos baratos y simples.

## 6 Agradecimientos

Agradecemos a Dra. Petra Wiederhold (CINVESTAV, México D.F.) por la revisión y corrección de este trabajo, sobre todo con respecto a la redacción en el idioma español. A los Dr. Herbert Suesse y Dr. Wolfgang Ortman agradecemos por su ayuda en el desarrollo den método SDR. También damos nuestras gracias a los revisores por sus comentarios valuales.

**Gamba, P., A.Mecocci, U.Salvatore.** "Vanishing point detection by a voting scheme". *Proc. ICIP Vol II, 1996, pp. 301-304*

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

**Hemmleb, M., A.Wiedemann.** "Digital rectification and generation of orthoimages in architectural photogrammetry". *Proc. CIPA, IAPRS Vol.32 part 5C1B (1997), pp. 261-267*

**Kanatani, K.** "Constraints on length and angle". *CVGIP* 4 (1988) 28-42

**Karras, G., G.Mountrakis, P.Patias, E.Petsa.** "Modelling distortion of super-wide-angle lenses for architectural and archaeological applications". *IAPRS, Vol.32, part 5 (1998) pp. 570-573*

**Liebowitz, D., A.Zisserman.** "Metric rectification for perspective images". *Proc. CCVPR 1998, pp. 482-488*

**Lutton, E., H.Maitre, J. Lopez-Krahe.** "Contribution to the determination of vanishing points using Hough transform". *IEEE Trans. PAMI* 16 (1994) 430-438

**McLean,G., D.Kotturi.** "Vanishing point detection by line clustering". IEEE Trans. PAMI 17 (1995) 1090-1095

**Y.Nomura,Y., M.Sagara, H.Naruse, A.Ide.** A simple calibration algorithm for high-distortion-lens camera. IEEE Trans. PAMI 14 (1992) 1095-1099

**Palmer,P.L., A.T.Tai.** "An optimised vanishing point detector". Proc. BMVC 1993, pp. 529-538

**Pratt,W.K.** "Digital Image Processing". John Wiley, New York 1991

**Prescott, B., G.McLean.** "Line-based correction of radial lens distortion". GMIP 59 (1997) 39-47

**Sawhney,H.S., R.Kumar.** "True multi-image alignment and its application to mosaicing and lens distortion correction". IEEE Trans. PAMI 21 (1999) 235-243

**Shih,S., Y.Hung, W.Lin.** When should we consider lens distortion in camera calibration. Pattern Recognition 28 (1995) 447-461

**Shah,S., J.K.Aggarwal.** "Intrinsic parameter calibration procedure for a (high distortion) fish-eye lens camera with distortion model and accuracy estimation". Pattern Recognition 29 (1996) 1775-1788

**Shufelt,J.A.** "Performance evaluation and analysis of vanishing point detection techniques". IEEE Trans. PAMI 21 (1999) 282-288

**Stein,G.P.** "Lens distortion calibration using point correspondences". Proc CVPR 1997, pp. 602-6

**Suesse,H. K.Voss, W.Ortmann, T.Baumbach.** Shift detection by restoration. Proc. 8th Conf. CAIP 1999, Ljubljana/Slovenia, Sept. 1999, pp.33-40

**Tsai,R.** "An efficient and accurate camera calibration technique for 3-D machine vision". IEEE Proc CCVPR 1986, pp. 364-74

**Van den Heuvel,F.** "Vanishing point detection for architectural photogrammetry". IAPRS, Vol.32, part 5, 1998, pp. 652-659

**Van den Heuvel,F.** "3D reconstruction from a single image using geometric constraints". ISPRS 53, 1998, pp. 354-368

**Voss,K., R.Neubauer, M.Schubert.** "Monokulare Rekonstruktion für Robotvision". Shaker-Verlag Aachen 1995, pp. 24

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

**Voss,K., W.Ortmann, H.SüBe.** "Apareamiento afín de parejas de imágenes". Proc. IV. Simposio Iberoamericano de Reconocimiento de Patrones, SIARP'99, Habana/Cuba, Marzo 1999, pp. 35-44

**Voss,K., H.Suesse, W.Ortmann, T.Baumbach** "Shift detection by restoration". Pattern Recognition 32 (1999) 2067-2068

**Weng,J., P.Cohen, M.Herniou.** "Camera calibration with distortion models and accuracy evaluation". IEEE Trans. PAMI 14 (1992) 965-980

**Weisstein,E.** "CRC Concise Encyclopedia of Mathematics" CRC Press, 1999, pp.1286

**WWW\_1 Voss,K., H.Suesse, J.L.Marroquín, S.Gutiérrez.** "Fundamentos y Nuevos Métodos para la Reconstrucción Tridimensional". [http://www.cimat.mx/proy\\_comp/libror3d/libror3d.html](http://www.cimat.mx/proy_comp/libror3d/libror3d.html), capítulo 8: [http://www.cimat.mx/proy\\_comp/libror3d/Postscript/cap8.ps](http://www.cimat.mx/proy_comp/libror3d/Postscript/cap8.ps) (1998)



**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 en 1969 en la misma universidad. Ha trabajado como profesor asistente en el Depto. de física y más tarde en el Depto. de informática. Desde 1971 ha trabajado en el área del procesamiento digital de imágenes, como ejemplo está su participación durante muchos años en el Depto. de Patología de la Universidad Humboldt, Berlin. En 1986 fue catedrático y jefe de un nuevo grupo de Procesamiento de Imágenes en la Universidad Friedrich Schiller (Alemania) en Jena, posición que actualmente conserva. Ha publicado más de 250 artículos científicos y siete libros sobre aspectos teóricos y prácticos del procesamiento de imágenes. Sus principales áreas de interés son la naturaleza discreta de las imágenes digitales, software interactiva para el procesamiento de imágenes y la reconstrucción 3D en la visión robótica.



**Christian Bräuer-Burchardt** estudio matemáticas en la Universidad Friedrich Schiller (Alemania) en 1987-1992. Ha trabajado como profesor asistente en el grupo Circulación retinal en el Depto. Oftalmología de la misma universidad, en este grupo ha desarrollado algoritmos y programas para la investigación de la circulación de la sangre por las venas y las arterias del ojo. Más tarde trabajó en proyectos de desarrollo de investigación en la Universidad Técnica Ilmenau (Alemania). Desde 1998 es profesor asistente en el grupo Procesamiento de Imágenes en la Universidad Friedrich Schiller. Su trabajo principal fue la elaboración de algoritmos y programas para la reconstrucción 3D de superficies grandes relacionadas con objetos arquitectónicos y con el movimiento de robots móviles. En ese tiempo escribió su tesis doctoral sobre reconstrucción monocular para objetos poliédricos con ayuda de información a priori. Ha publicado 15 artículos científicos.

