

ESTIMACIÓN DE LA ALTURA DE LOS EDIFICIOS MEDIANTE TRATAMIENTO DIGITAL DE IMÁGENES

REFERENCIA PACS: 43.50.Sr

García Duro, José María; Nava Baro Enrique

¹Departamento de Ingeniería de Comunicaciones. ETSI de Telecomunicación.

Campus de Teatinos, Universidad de Málaga. 29071 Málaga (España).

email: josemariagarciauro@gmail.com; en@uma.es;

ABSTRACT

Noise mapping estimation needs accurate building models. This paper demonstrates the use of epipolar geometry to determine building geometry parameters (i.e. height) using digital photographs of buildings that can be obtained with a camera mounted in a vehicle. Focusing on well-known geometrical methods, a reconstruction tool is implemented which use pairs of stereo photographs as source. A calibrated image processing algorithm has been developed which employes different methods for image matching and Ransac estimation of the fundamental matrix, Geometric models with limited error range are finally obtained.

RESUMEN

La realización de mapas de ruido requiere un modelado de los edificios precisa. En este trabajo se presenta el uso de la geometría epipolar para estimar parámetros geométricos de los edificios (por ejemplo, la altura) utilizando fotografías digitales de los mismos que pueden obtenerse con una cámara instalada en un vehículo. Utilizando métodos geométricos conocidos se ha implementado una herramienta de reconstrucción que utiliza pares de fotografías estéreo como entrada. Se ha desarrollado un algoritmo de tratamiento digital de imágenes que utiliza diferentes técnicas para correlacionar ambas imágenes estéreo y realizar una estimación Ransac de la matriz fundamental. Con este algoritmo se obtienen modelos geométricos de los edificios con un error controlado.

INTRODUCCIÓN.

El objetivo de este trabajo es desarrollar un método para obtener medidas topográficas de edificios mediante el uso del análisis de imágenes y el tratamiento digital de las mismas, con el objetivo básico de obtener datos con un nivel de precisión adecuado y un coste en tiempo y dinero muy inferior al de las técnicas topográficas clásicas.

La idea básica sobre la que se desarrollará la aplicación será obtener a partir del procesado de una serie de fotografías de un edificio, unas medidas de altura, anchura del mismo, de manera que con esos datos y los datos propios sobre las fuentes de ruido podamos obtener mapas de ruido mucho más realistas y sobre todo mucho más eficientes en cuanto a la inversión económica necesaria en su realización.

El objeto es por tanto el estudio de las técnicas actuales de reconstrucción de edificios, y la implementación de un método que se apoye en las ideas estudiadas para obtener una reconstrucción con un nivel de precisión razonable, empleando geometría epipolar [1].

DESCRIPCIÓN DEL ALGORITMO PROPUESTO.

Geometría epipolar

La geometría epipolar es la geometría proyectiva intrínseca que se da entre dos vistas. Es independiente de la estructura de la escena y sólo depende de los parámetros internos de las cámaras y las posiciones relativas entre las mismas.

Básicamente, es la geometría que se da cuando se intersectan los planos imagen con el conjunto de planos, que tienen como eje, la línea que une los centros de las cámaras. Esta geometría generalmente aparece cuando se buscan correspondencias de puntos entre un par de imágenes estereó, para entenderla se comenzará desarrollando este concepto. Supóngase un punto X en un espacio tridimensional que puede visualizarse en dos proyecciones. El problema que se pretende resolver es: ¿qué relación existirá entre x y x' ? los puntos correspondientes en cada imagen. En la figura 1 se presenta la definición del plano epipolar π , con la correspondencia de los puntos x, x' con el punto 3d X y los centros de las cámaras C y C' , que son coplanares.

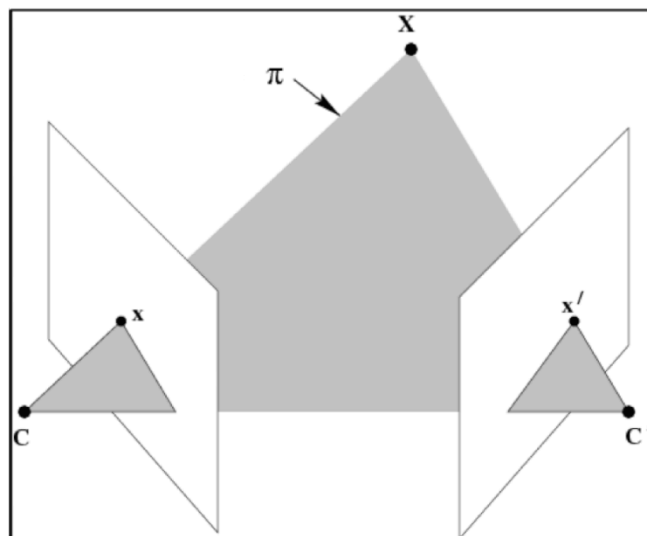


Figura 1. Geometría de la correspondencia entre x y x' .

La correspondencia entre las dos proyecciones vendrá dada por la matriz fundamental F , que relaciona los puntos entre diferentes proyecciones mediante la ecuación

$$x F x' = 0,$$

donde F es una matriz 3×3 . Debe tenerse en cuenta que como F tiene siete grados de libertad, es necesario disponer de al menos 7 puntos emparejados para definir el plano epipolar. La ventaja fundamental del uso de la matriz fundamental es su independencia de la calibración de la cámara, ya que sólo depende de las proyecciones.

Estimación de la matriz fundamental

Para el cálculo de la matriz fundamental existen diferentes algoritmos, aunque como en este trabajo se pueden obtener más de 7 puntos emparejados, se utilizará el método de Levenberg-Marquard [2] en la minimización de la distancia geométrica.

Para la obtención de los puntos notables que permiten obtener la matriz fundamental F se utilizará una versión modificada de un algoritmo de tratamiento digital de imágenes llamado detector de Harris [3], que fue inicialmente diseñado para la detección de puntos singulares, como por ejemplo, las esquinas. La modificación realizada consiste en la aplicación de un filtro paso bajo no lineal para la reducción de las texturas de la imagen, que tienen un efecto desfavorable sobre la estimación de la geometría epipolar.

Una vez obtenidos los puntos notables se procede a un emparejamiento, en el que se han evaluado dos posibilidades: un algoritmo clásico de correlación restringido a una zona de búsqueda y un algoritmo de información local en frecuencia [4], que se basa en el uso de la información en fase de un punto de una señal para guiar la búsqueda del punto que en otra imagen más le corresponda, utilizando filtros de Gabor en cuadratura para obtener la información local de fase y frecuencia instantánea. La disparidad se calcula al dividir la fase local entre la frecuencia instantánea siendo ésta última calculada a partir de la derivada de la fase local, empleando una transformada wavelet usando un banco de filtros creados a partir del escalamiento geométrico de una forma de onda, de manera que cada filtro captura una banda determinada de frecuencias, y que estas bandas se relacionan de manera logarítmica entre ellas. Una estrategia similar se usa para estimar la información local de fase.

Una vez calculadas las matrices de calibración de la cámara empleada en la adquisición de las imágenes, puede obtenerse la geometría epipolar con la obtención del epipolo y las líneas epipolares, a partir de las cuales se pueden estimar las alturas de los edificios y, en general, un modelo 3D completo de los mismos.

En el siguiente apartado se va a realizar una descripción de la aplicación del método propuesto en el caso de un edificio cuyas dimensiones se han comprobado por medida directa. Se presentarán los resultados obtenidos tras cada una de las diferentes fases del algoritmo y una breve discusión de los mismos.

RESULTADOS

En la Figura 2 se presenta un ejemplo de los resultados obtenidos. Se observa las dos proyecciones de un edificio de una planta, tomadas con la misma cámara a la misma altura, pero con la segunda proyección en una posición desplazada a la derecha a una distancia respecto a la anterior posición de la cámara conocida. En las imágenes se han superpuesto dos polígonos en color rojo, que muestran la zona de interés que es común a ambas imágenes, otros dos de color azul, que indican la zona de texturas de origen natural y otras dos de color verde, con las texturas artificiales, que serán de utilidad para describir los resultados obtenidos.



Figura 2. Edificio de una planta. Dos proyecciones.

En este trabajo se ha elegido este ejemplo, porque las imágenes presentadas son casos de gran dificultad para la reconstrucción automática, debido a la presencia en las imágenes de zonas con texturas muy definidas, como el tejado de la casa y las copas de los árboles cercanos. Ello es así porque el detector de Harris, utilizado para la detección automática de los puntos emparejados, es muy sensible a las texturas, lo que provocará un gran número de puntos notables, como puede observarse en la Figura 3, donde se obtiene un total de 3552 puntos en la imagen correspondiente a la proyección izquierda y 2921 en la imagen de la proyección derecha. Puede observarse que las zonas con texturas finas conducen a un gran número de puntos, así como existen gran número de puntos fuera de las zonas de las imágenes que son comunes, como las sobreimpresas en la Figura 2. Sin la aparición de dichas texturas, los puntos notables coincidirían con vértices y puntos singulares de los edificios, lo que mejoraría las prestaciones del algoritmo propuesto. En este caso, se ha realizado un filtrado de mediana para reducir el efecto de la textura, pero el objetivo futuro de trabajo sería la detección de las texturas habituales en un ambiente urbano para su eliminación previa a la aplicación del detector de Harris.



Figura 3. Puntos notables detectados en las dos proyecciones.

La tercera fase del algoritmo consiste en la aplicación de un algoritmo para determinar el emparejamiento probable de los puntos notables que el algoritmo de Harris ha obtenido. Se han implementado dos algoritmos diferentes: un algoritmo de correlación y otro de información local en frecuencia. En la Figura 4 se muestra una imagen que contiene las dos proyecciones superpuestas con una capa de transparencia del 50% mostrando los resultados obtenidos con el método de correlación. Sobre la imagen se muestran los vectores resultado del algoritmo de

correlación empleado, restringido a un radio de búsqueda de 200 píxeles y con un tamaño de ventana de 21x21. Según puede apreciarse, la zona marcada en la Figura 2 con un polígono de color verde, que corresponde al tejado de la vivienda, zona con una textura de tipo regular, produce un gran número de emparejamientos entre los puntos notables de ambas proyecciones, de lo que se desprende que este tipo de texturas no tienen un efecto negativo sobre la estimación de la matriz fundamental F. Por el contrario, los puntos situados en la zona azul tienen un comportamiento más errático en el emparejamiento, ya que provienen de texturas de origen natural, algunos de estos emparejamientos son claramente incorrectos y dificultarán la estimación de la matriz F. En los puntos de la zona roja, que corresponde a la zona sin texturas definidas, en general, se obtienen buenas correspondencias, aunque algunas de ellas son incorrectas. En la parte inferior de la figura se muestran algunos ejemplos de los resultados obtenidos. El número de puntos emparejados en este ejemplo es de 850, un 29% del total de puntos notables detectados en la fase anterior.



Figura 4. Resultados del emparejamiento por correlación. En las subimágenes inferiores, detalles de la imagen superior.

En la figura 5 se muestran los resultados obtenidos con el método de información local en frecuencia. En este caso puede apreciarse como las texturas regulares (tejado) aportan menos información al emparejamiento, así como en el caso de las texturas de origen natural (copa del árbol), pero en la zona que contiene las líneas maestras del edificio, la relación entre emparejamientos correctos e incorrectos mejora sensiblemente. La conclusión obvia es que este algoritmo funciona mejor con imágenes que contienen áreas importantes con texturas. En términos cuantitativos, se han obtenido 846 puntos emparejados, un número casi idéntico al obtenido con el método de correlación.

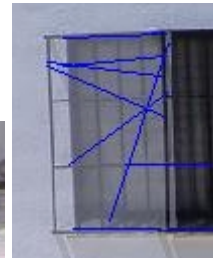
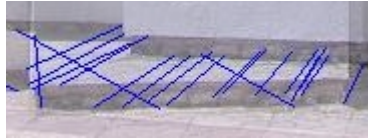
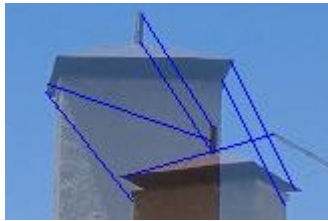
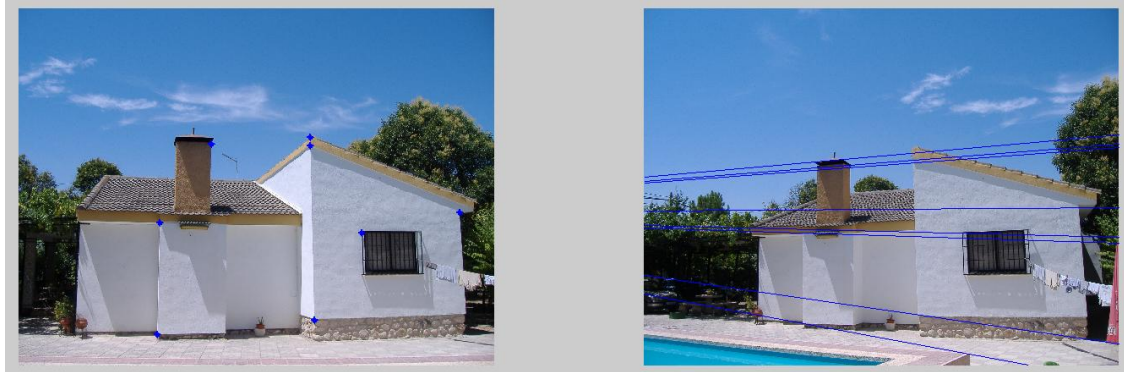


Figura 5. Resultados del emparejamiento por información local en frecuencia. En las subimágenes inferiores, detalles de la imagen superior.

A continuación, con los emparejamientos obtenidos se estima la matriz fundamental F que, junto con la matriz de calibración de la cámara, nos permite obtener la geometría epipolar. En la figura 6 se presentan los resultados obtenidos con las dos técnicas de emparejamiento utilizadas, la de correlación y la de información local en frecuencia. En las imágenes de la izquierda se han marcado los puntos singulares detectados que pueden utilizarse para la estimación de las dimensiones del edificio, mientras que en las imágenes de la derecha se han resaltado las líneas epipolares estimadas, que coinciden en un punto llamado epipolo. Puede observarse que con ambos métodos la geometría epipolar obtenida es bastante similar, si bien en el método por información local en frecuencia se observa que el epipolo queda también fuera de la imagen pero más cercano, lo que indica que con este método las cámaras serán restituidas más próximas una a la otra.

Correlación



Información local en frecuencia

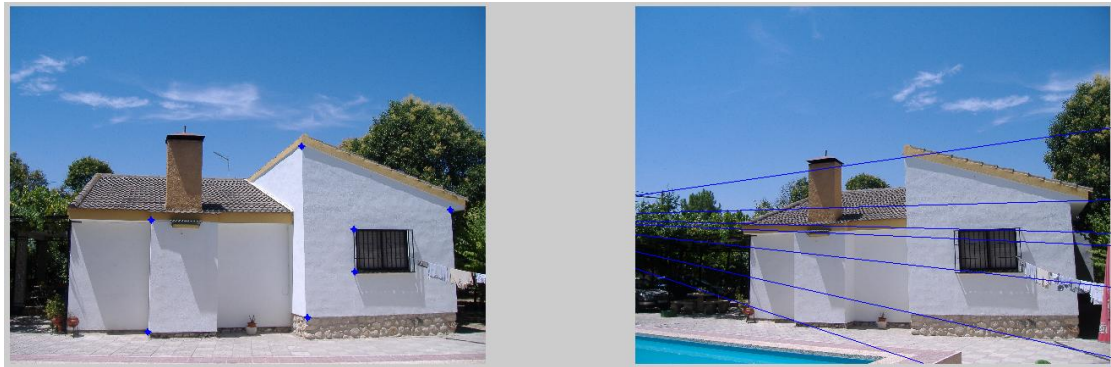


Figura 6. Geometría epipolar. En la parte superior se presenta los resultados obtenidos por correlación y en la parte inferior mediante el método de información local en frecuencia.

Finalmente, y con la ayuda de la referencia de distancia entre las posiciones de la cámara en la obtención de las dos imágenes, se obtienen las dimensiones del modelo 3D del edificio reconstruido. En la figura 7 se presentan los resultados obtenidos en cuanto al error porcentual de las estimaciones con respecto a las medidas reales del edificio empleando una referencia vertical. En este caso, los errores suelen ser muy reducidos en todas las dimensiones verticales, en torno al 1-2%, mientras que aumentan considerablemente en la dimensión horizontal, típicamente 10-15%, llegando en un caso al 31%. En el caso de utilizar una referencia horizontal, se obtendría una inversión de los errores, siendo mayores en los elementos verticales y menores en los horizontales. Esto es debido a que el propio método de triangulación utilizado para estimar los puntos 3D, es muy dependiente de la geometría epipolar, y más concretamente de la matriz fundamental, que no tendrá el mismo comportamiento en el eje vertical que en el eje horizontal, ya que obviamente la reconstrucción no es perfectamente euclídea, y las direcciones que en el espacio euclídeo real son perpendiculares no serán perfectamente perpendiculares en la reconstrucción. Una manera de solucionar este problema sería tener dos referencias, una en cada eje, y combinar ambas reconstrucciones, aunque para ello tendríamos que extender el método y el número de proyecciones, lo que en un ambiente de elaboración de mapas de ruido no parece excesivamente complicado.



Figura 7. Error porcentual en las medidas, empleando una referencia vertical.

CONCLUSIONES

La contaminación acústica hoy en día es un problema en todas las ciudades, las leyes actuales en la materia obligan al trazado de mapas acústicos, estos mapas acústicos necesitan los datos más actualizados de los alzados de los edificios para poder estimar mediante simuladores como afecta el ruido en determinadas zonas.

En este trabajo se ha abordado el problema de la reconstrucción de un modelo 3D de un edificio a partir de unos datos de entrada mínimos, con el objetivo final de poder usar esta reconstrucción como base para ayudar elaborar el trazado de un mapa de ruido en una sección urbana.

Se han estudiado varias opciones para la reconstrucción 3D de edificios, desde distintos enfoques y con distintos datos de entrada a los diferentes algoritmos, empleando geometría epipolar y fotografías de los edificios. Tras las simulaciones realizadas, se ha podido concluir que:

- Es esencial que el edificio a reconstruir tenga el máximo solape en ambas fotografías, fotografías muy dispares llevan a una reconstrucción pobre.
- Aunque en un principio las texturas podrían ser útiles para el proceso de emparejamiento de puntos notables, a la hora de detectar los mismos es beneficioso intentar eliminarlas en la medida de lo posible.
- El detector de esquinas de Harris, complementado con un filtro de mediana se muestra suficientemente potente para detectar suficientes puntos notables y con la suficiente singularidad para poder ser emparejados.
- Entre los dos métodos de emparejamiento analizados, el emparejamiento por información local en frecuencia es más costoso en tiempos de cálculo pero tiene un mejor comportamiento a la hora de detectar puntos notables en zonas con menos textura.
- La reconstrucción obtenida es sensible a la medida que se tome como referencia, pero en general se obtienen resultados que varían entre un 1% y un 10%. Este error

puede reducirse en un futuro, cuando se disponga de una base más amplia de imágenes para optimizar los algoritmos implementados.

Los resultados obtenidos indican que con el método propuesto se puede obtener una estimación razonable de las alturas de los edificios partiendo de dos fotografías tomadas con una cámara y una referencia adicional, tal como las coordenadas geométricas de las posiciones de la cámara en la que se obtuvieron las fotografías, información que se puede obtener mediante un dispositivo GPS. Además de las alturas de los edificios, puede obtenerse también una descripción geométrica completa de los edificios, que podría ser exportada en formatos de uso común, tales como el estándar VRML.

Aunque en este trabajo sólo se han presentado resultados preliminares de la aplicación del algoritmo propuesto, es necesario una evaluación más completa con un mayor número de casos.

REFERENCIAS

- [1] Xu,G. y Zhang,Z. *Epipolar Geometry in Stereo, Motion and Object Recognition*. Kluwer Academic Publishers, 1996.
- [2] Press,W. Flannery,B. Teukolsky,S. y Vetterling,W. *Numerical Recipes in C*. Cambridge University Press, 1988..
- [3] Harris,C. y Stephens,M.J. "A combined corner and edge detector". In Alvey Vision Conference, Manchester, pp. 147 - 152, 1988
- [4] Kovesi,P.D. "Image Correlation From Local Frequency Information" The Australian Pattern Recognition Society Conference: Dicta 95. pp 336-341. December 1995. Brisbane.