

# PANORAMAS

CONSTRUCCIÓN DE PANORAMAS LINEALES CON PROYECCIÓN EN SUPERFICIES CILÍNDRICAS  
Y MEZCLA DE IMÁGENES USANDO EL ALGORITMO DE BURT-ADELSON

## CONTENIDO

Introducción.....	3
Enfoque.....	3
Paso a cilíndricas.....	3
Selección de puntos.....	4
Cálculo homografías.....	6
Selección de áreas.....	7
Mezcla (pirámide).....	10
Pros y contras.....	14
Pros.....	14
Contras.....	15
Posibles propuestas de mejora.....	16
Resultados finales.....	17
Manual de usuario.....	17
Referencias.....	17

## INTRODUCCIÓN

La construcción de panoramas a partir de imágenes individuales siempre ha tenido gran interés tanto comercial como personal. A la hora de realizar una fotografía de una escena, los fotógrafos se han visto limitados por las características de la cámara y han tenido que renunciar a gran parte de la escena. Si estos querían realizar una imagen panorámica debían recurrir a costosos artilugios o a la tediosa labor de unir varias fotografías a mano.

Gracias a los avances de la tecnología, ahora es posible la construcción de panoramas de forma automática, lo que supone una gran ayuda y habilita a todo fotógrafo a construir de manera simple y sencilla panoramas de calidad.

## ENFOQUE

De forma general, el proceso que se ha seguido para construir los panoramas ha sido el siguiente:

- Transformación a proyección cilíndrica de las imágenes.
- Cálculo de los puntos destacados con SIFT.
- Cálculo de las parejas en correspondencias con el método K-NN y el criterio de Lowe.
- Estimación de las homografías entre cada par de imágenes.
- Cálculo de las áreas que se van a mezclar.
- Utilización del algoritmo de mezcla piramidal con ambas áreas.
- Unión de las imágenes en un único panorama.

## PASO A CILÍNDRICAS

El primer paso para la construcción del panorama consiste en pasar la imagen a una proyección en superficie cilíndrica. Podemos hacerlo haciendo una transformación hacia delante (Forward Warping) o haciendo una transformación inversa (Inverse Warping), el primero es más sencillo y consiste en calcular el desplazamiento de cada pixel de la imagen original a la imagen transformada. Pero se corre el riesgo de que queden regiones sin colorear, por tanto se ha utilizado la transformación inversa que consiste en buscar el color que le corresponde a cada pixel de la imagen transformada desde la imagen original.

Sea  $(x, y)$  un punto de la imagen original y  $(x', y')$  su correspondencia en la imagen transformada, para calcular a qué  $(x, y)$  le corresponde un  $(x', y')$  dado, aplicamos:

$$x = f \cdot \tan\left(\frac{x'}{s}\right)$$
$$y = \frac{y' \cdot f \cdot \sqrt{1 + \tan^2\left(\frac{x'}{s}\right)}}{s}$$

Donde  $f$  es la distancia focal de la imagen y  $s$  un factor de escala. En nuestro caso, usaremos  $s=f$ , por lo que las fórmulas utilizadas son:

$$x = f \cdot \tan\left(\frac{x'}{f}\right)$$
$$y = y' \cdot \sqrt{1 + \tan^2\left(\frac{x'}{f}\right)}$$

El problema de la transformación cilíndrica es que, salvo que se conozca la calibración de las cámaras, no es posible determinar qué valor usar para el factor de escala. Por tanto, para cada grupo de imágenes, se ha de recurrir a ensayo y error para determinar el valor apropiado del factor de escala. En las imágenes siguientes, se puede apreciar la diferencia de usar distinto valor.

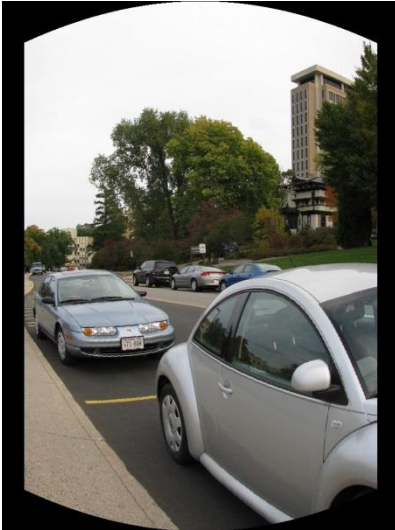


Fig. 1  $f=600$

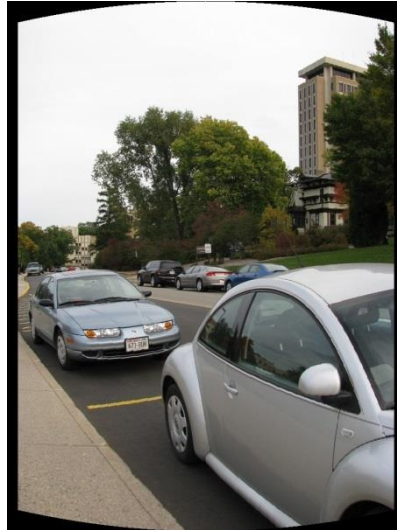


Fig. 2  $f=900$

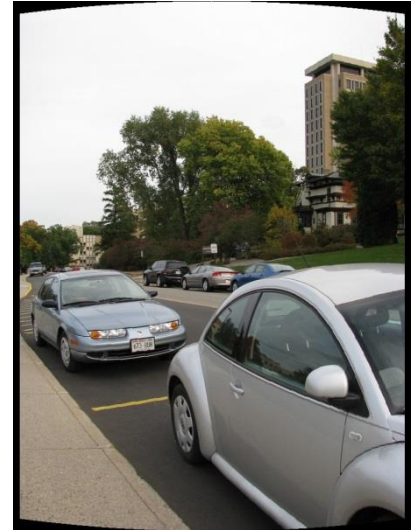


Fig. 3  $f=1200$

#### SELECCIÓN DE PUNTOS

A continuación hemos de calcular las correspondencias entre cada par de imágenes, para ello, primero se deben buscar los puntos destacados de cada imagen. En nuestro caso, hemos optado por usar el algoritmo SIFT [Fig. 4] con los parámetros indicados en la Tabla 1. Estos valores han sido fijados tras realizar varias pruebas y observar los resultados.

Parámetro	Valor
nFeatures	0
nOctaveLayers	3
contrThres	0,05
edgeThres	10
sigma	1,6

Tabla 1 Parámetros SIFT

De todos los parámetros disponibles en la implementación de OpenCV para SIFT, se han dejado por defecto, por ofrecer buenos resultados, los parámetros siguientes:

- El valor de **sigma** usado para aplicar el filtro gaussiano al bajar niveles en la pirámide.
- El nº máximo de puntos a retener (**nFeatures**), que en este caso, al ser 0 significa que no vamos a descartar ningún punto encontrado.
- El umbral para bordes (**edgeThres**). Reducir este valor significaba perder muchos puntos destacados mientras que subirlo por encima de 10 apenas reportaba nueva información.
- El número de capas intermedias entre cada nivel de la pirámide (**nOctaveLayers**). Se mantiene en 3 por no apreciarse cambios al variar el valor.

En cambio, sí se ha modificado el parámetro que controla el umbral de contraste (**contrThres**). Se ha aumentado un poco para reducir el número de puntos destacados ya que con un valor inferior se obtenían muchos puntos que estaban muy cercanos entre sí.

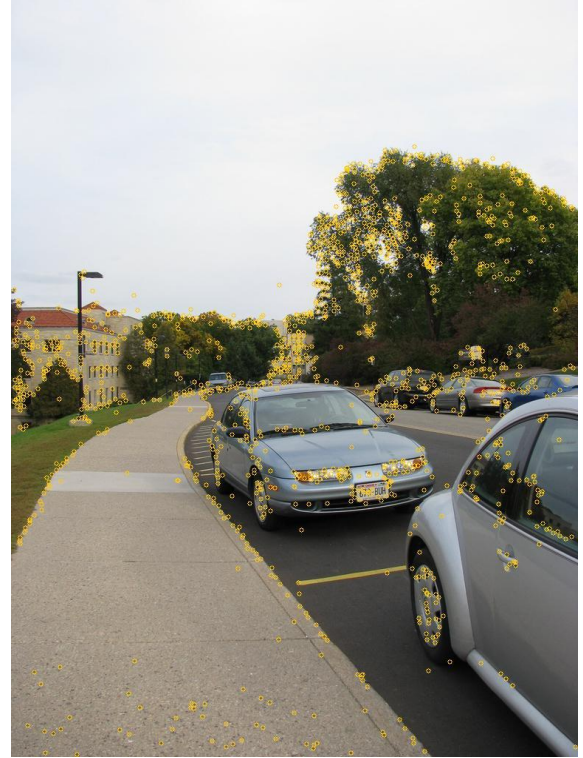


Fig. 4 Puntos SIFT destacados en amarillo

Con los puntos SIFT extraídos de cada imagen, el siguiente paso ha sido establecer las correspondencias realmente buenas entre cada pareja de imágenes. En este caso, la técnica usada ha sido la búsqueda K-NN junto al criterio de Lowe con  $k=2$  y  $\text{umbral}=0,4$  [Fig. 5].



Fig. 5 Correspondencias encontradas con K-NN y método Lowe

Esta técnica consiste en buscar, para cada descriptor  $x$  de la primera imagen, sus dos mejores correspondencias  $f_1$  y  $f_2$  en la segunda imagen. Después, se divide la distancia de  $x$  y  $f_1$  entre la distancia de  $x$  y  $f_2$ . Se aceptará la correspondencia  $(x, f_1)$  sólo si no se supera el umbral definido. Es decir:

$$\frac{\|x - f_1\|}{\|x - f_2\|} < \text{umbral}$$

Para definir el umbral se ha de saber que valores grandes denotan correspondencias ambiguas y valores pequeños correspondencias certeras. Así reducimos la cantidad de puntos iniciales a un pequeño grupo de parejas correctas.

#### CÁLCULO HOMOGRAFÍAS

Con las parejas en correspondencias calculadas, podemos estimar las homografías para unir las imágenes. Se ha escogido el método RANSAC por su buen comportamiento al tratar con outliers. El procedimiento que hemos seguido ha sido el que aparece esquematizado en la figura 6 y se basa en dos puntos:

- Conocemos la situación de la imagen central puesto que la queremos situar en el centro del canvas.
- Haciendo uso de la composición de las homografías, calculamos dónde irá situada la siguiente imagen a partir de la homografía de la imagen anterior multiplicada por la suya propia, es decir:  $H_x = H_{x-1} \cdot H_x$

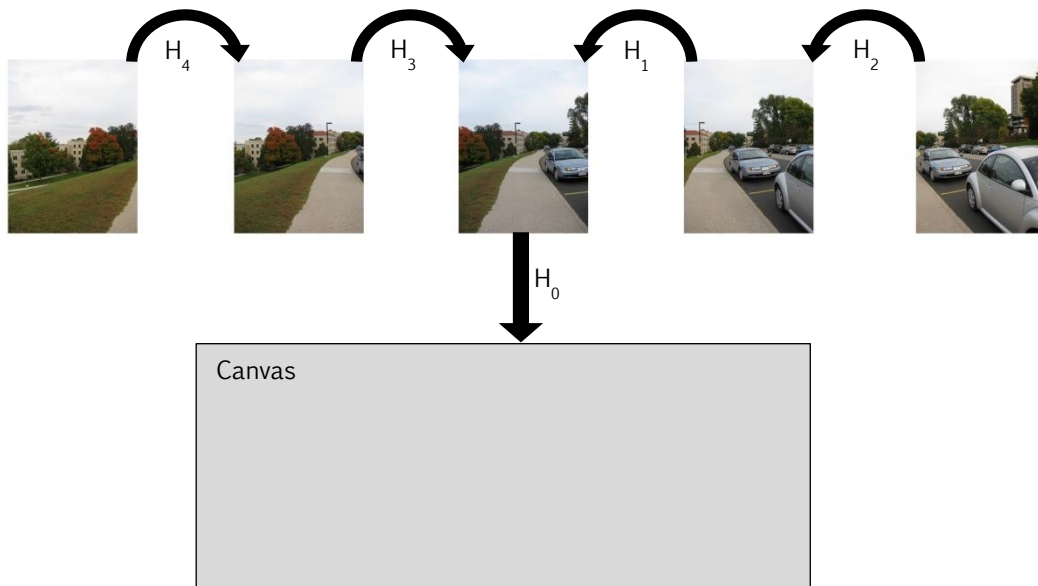


Fig. 6 Esquema para el cálculo de las homografías



Por tanto, según el esquema anterior, comenzamos tomando la imagen central y asignándole una homografía prestablecida que la traslada aproximadamente al centro del canvas:

$$H_0 = \begin{pmatrix} 1 & 0 & \frac{Largo\_Canvas}{2} \\ 0 & 1 & \frac{Alto\_Canvas - Cols\_Imagen}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

Después se calculan las homografías de las imágenes a la derecha de la central, siguiendo el sentido de izquierda a derecha ( $H_1$  y  $H_2$  en la figura) y componiendo cada una con la anterior. Y para acabar se repite la operación para las imágenes a la izquierda de la central pero esta vez, calculando las homografías de derecha a izquierda.

#### SELECCIÓN DE ÁREAS

En este punto tenemos todos los ingredientes para montar un panorama con una buena continuidad geométrica entre las imágenes unidas gracias a las transformaciones que nos pueden proporcionar las matrices de homografía calculadas. Sin embargo, es muy fácil que por cualquier motivo (cambio en la exposición o del balance de blancos, pasa una nube, un reflejo en la lente...) no haya continuidad de color en el punto de unión produciéndose un corte notorio en el panorama [Fig. 19 izquierda].

Para solventar el problema de la continuidad del color existen varias técnicas conocidas. Nosotros hemos optado por la mezcla piramidal ya que además de proporcionar unos resultados muy buenos con una transición suave, es bastante robusta de cara a cambios en la exposición entre las imágenes a mezclar [Fig. 19 derecha]. El procedimiento para realizar la mezcla piramidal se describe más adelante ya que antes es necesario determinar la región adecuada de ambas imágenes para aplicar la mezcla.

Tal y como se ha explicado en el cálculo de homografías, las transformaciones necesarias a aplicar sobre cada imagen del panorama se han calculado con respecto de la imagen central, cuya homografía consiste únicamente en una traslación. Pero, una vez conocida la homografía de cada imagen es más simple montar el panorama de izquierda a derecha de modo que en cada iteración la imagen de la izquierda (panorama realizado hasta el momento) se queda intacta y a la de la derecha se le aplica la homografía que le corresponda. Obviamente la primera imagen del panorama es necesario colocarla antes del bucle iterativo descrito.

En un montaje de panorama sin mezcla piramidal con ir colocando en cada iteración la imagen de la derecha donde corresponda solapándose sobre el panorama acumulado hasta el momento es suficiente.

Dado que nuestra intención es aplicarle la mezcla piramidal a la unión de ambas imágenes antes de proceder a solapar la de la derecha sobre la de la izquierda hay que determinar una región común entre ambas imágenes para seleccionar exactamente la misma parte del panorama correspondiente a las dos imágenes. Una vez conseguido este objetivo ya podemos aplicar la mezcla piramidal.

En este punto hay que distinguir entre la región común entre ambas imágenes [Fig. 7] y la región seleccionada que es aquella que usaremos para realizar la mezcla [Fig. 8]. La región común se puede calcular a partir de algunos puntos extremos de las imágenes. Para conocer dichos puntos basta con calcular la localización de esos puntos de las imágenes originales aplicándoles sus homografías correspondientes:

$$(x', y') = \left( \frac{h_{11}x + h_{12}y + h_{13}}{h_{31}x + h_{32}y + h_{33}}, \frac{h_{21}x + h_{22}y + h_{23}}{h_{31}x + h_{32}y + h_{33}} \right), \text{ donde } h_{xy} \text{ son las coordenadas de la homografía } H$$

En la figura 7 se aprecia una línea negra donde termina la imagen de la izquierda y comienza la de la derecha. Si no fuéramos a aplicar la mezcla piramidal sería necesario suprimir el borde negro de la imagen de la derecha para una buena continuidad geométrica. Como vamos a aplicar mezcla piramidal ese borde no afecta al resultado debido a que en el extremo izquierdo de la selección se impondrá el contenido de la imagen de la izquierda.



Fig. 7 Región común entre dos imágenes

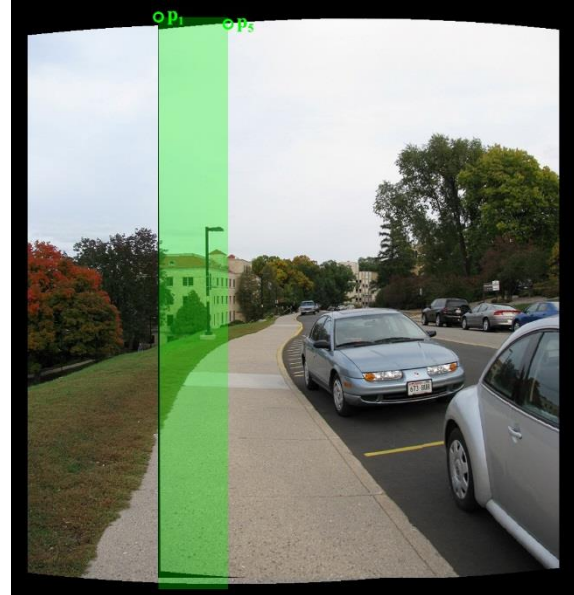


Fig. 8 Región seleccionada para mezclar

El punto  $p_1$  corresponde a la posición del punto (0,0) de la imagen de la derecha original aplicándole su correspondiente homografía. El punto  $p_2$  corresponde a la posición del punto (im\_or.cols-1,0) de la imagen de la izquierda original donde im\_or.cols es el ancho de dicha imagen. Considerando  $p_{1x}$  y  $p_{2x}$  la coordenada x de los puntos  $p_1$  y  $p_2$  respectivamente,  $\text{región\_común} = p_{2x} - p_{1x}$  es el ancho de la región que tienen en común ambas imágenes. El tamaño de esta región dependerá directamente del solapamiento que haya entre las imágenes de entrada. Si dichas imágenes están muy poco solapadas nos interesará aprovechar el máximo posible de la región común. Sin embargo, tal y como se puede apreciar en este ejemplo hay casos en los que la región solapada es muy grande con respecto al tamaño de la imagen. Si aprovechamos toda el área solapada acabaremos haciendo la mezcla piramidal sobre una selección innecesariamente grande. Es por ello que se ha optado por limitar el tamaño máximo del ancho de la región común al 20% del ancho de la imagen de entrada.

Por otra parte hay que tener en cuenta que para el correcto funcionamiento de la mezcla piramidal hay que seleccionar una región de un tamaño tal que podamos asegurar que las filas y las columnas sean divisibles entre 2 las veces necesarias para aplicar la mezcla piramidal. El número de filas de la selección no es necesario controlarlo dado que se mantiene la altura de las imágenes de entrada que se suponen de las dimensiones adecuadas. Además, en función de dicha altura se determina el número máximo de niveles que pueden aplicarse en la mezcla piramidal. Éste valor máximo obtenido se usará como límite si en algún momento el número de columnas de la selección permiten aplicar más niveles en la mezcla piramidal. En caso de que se determine un valor máximo de niveles superior a 5 éste se limita a 5 debido a que no se ha notado mejora en el resultado de la mezcla piramidal por encima de 5.



Dado que en el ejemplo seguido para esta explicación la región común es mucho mayor que el 20% del ancho la imagen de la derecha procedemos a limitar el tamaño de dicha región. El punto  $p_4$  nos muestra donde queda nuestro extremo derecho de la región común a considerar. Ahora ya podemos aplicar la mezcla piramidal sobre una selección del alto completo de la región solapada y de ancho  $= p_{4x} - p_{1x}$ . No obstante, en este momento no se puede asegurar que el valor de ancho sea divisible entre dos las suficientes veces como para que se puedan aplicar los niveles de pirámide necesarios para obtener una buena mezcla. Por este motivo se procede a disminuir el valor de ancho. El criterio de disminución del ancho puede tener dos criterios de parada: o ya es posible de dividir entre 2 tantas veces como el valor máximo calculado para las filas inicialmente o bien no se ha alcanzado dicho valor en cuyo caso nos quedamos con el que más veces divisible entre 2 es. En cualquier caso no bajamos de la cuarta parte del ancho de la región limitada.

En la figura 8 podemos ver la región común final seleccionada para aplicar la mezcla piramidal. Consiste en un rectángulo con esquina en  $p_1$  de ancho  $p_{5x} - p_{1x}$  y de la altura completa de la imagen de entrada.

En las figuras 9 y 10 se puede ver la región seleccionada sobre la imagen de la izquierda y la de la derecha por separado. Se puede observar que la imagen de la izquierda tiene una gran parte sin seleccionar a la derecha de la región a mezclar. Esto no tiene importancia debido a que la imagen de la derecha va a superponerse sobre esa región.

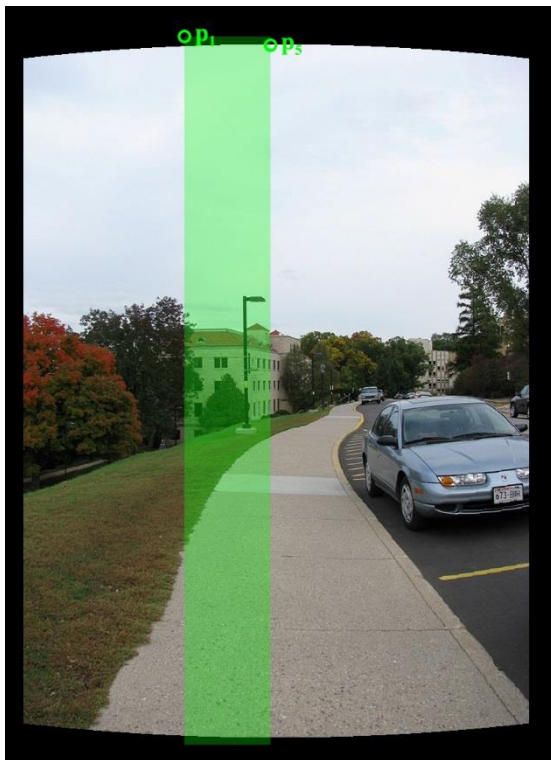


Fig. 9 Región seleccionada en la imagen izquierda



Fig. 10 Región seleccionada en la imagen derecha

Una vez calculada la mezcla piramidal que se describirá en el apartado siguiente, la región seleccionada de la imagen de la derecha la sustituimos directamente por el resultado obtenido de la mezcla piramidal. Ahora, para superponer la imagen de la derecha sobre la de la izquierda en la posición adecuada ya disponemos de la esquina superior izquierda ( $p_1$ ), nos falta calcular la posición de la esquina inferior derecha a partir de la imagen original aplicándole la homografía correspondiente. Este punto lo llamaremos  $p_3$ . En la figura 11 podemos ver sobre la

imagen de la izquierda la región en la que debe colocarse la imagen de la derecha. En la figura 12 vemos la misma región seleccionando la imagen de la derecha.



Fig. 11 Región donde se colocará la imagen de la derecha



Fig. 12 Región donde se colocará la imagen de la izquierda

#### MEZCLA (PIRÁMIDE)

La mezcla piramidal nos permite obtener transiciones suaves entre imágenes gracias a la recomposición de las imágenes a distintas frecuencias tal y como viene descrito en el artículo de Burt y Adelson. Ciñéndonos al uso de las pirámides gaussianas para obtener las laplacianas y combinar las imágenes a mezclar en distintas frecuencias, para un correcto funcionamiento del algoritmo de mezcla piramidal es necesario que la región de entrada seleccionada tenga una dimensión adecuada para que al subir y bajar por los niveles de la pirámide ningún redondeo cause que dejen de coincidir las dimensiones de las imágenes a mezclar. Esta condición y la forma de satisfacerla viene ampliamente descrita en el apartado anterior.

Vamos a mostrar el proceso de la mezcla piramidal sobre el mismo ejemplo que el seguido en el apartado anterior. Para compactar la representación no se va a mostrar toda la altura de las imágenes intermedias del proceso si no que únicamente la cuarta parte de la altura mostrando la parte central.

En primer lugar es necesario calcular la pirámide gaussiana de tantos niveles como vayamos a emplear en el proceso de mezcla. El nivel cero es la imagen original. Para la obtención de los niveles siguientes en primer lugar se aplica un filtro de convolución gaussiano y posteriormente se realiza un submuestreo a la mitad de filas y de columnas. En las figuras 13 y 14 podemos ver la pirámide gaussiana correspondiente a las imágenes de entrada por separado.

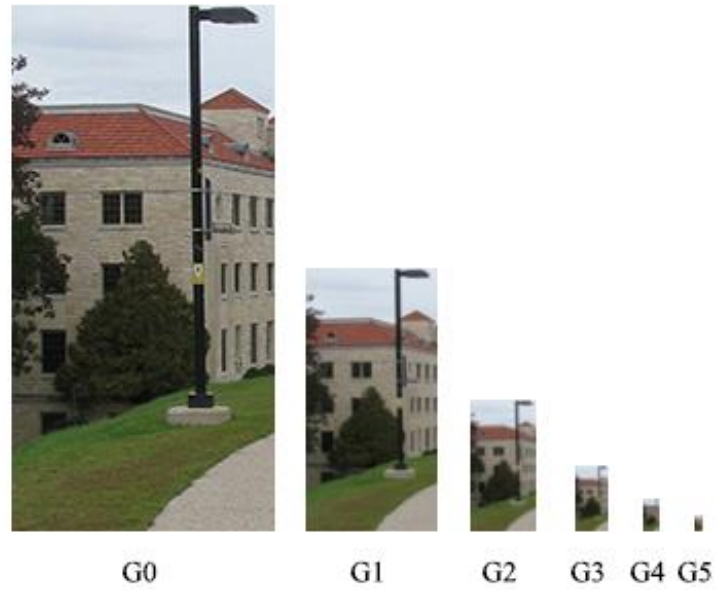


Fig. 13 Pirámide gaussiana izquierda

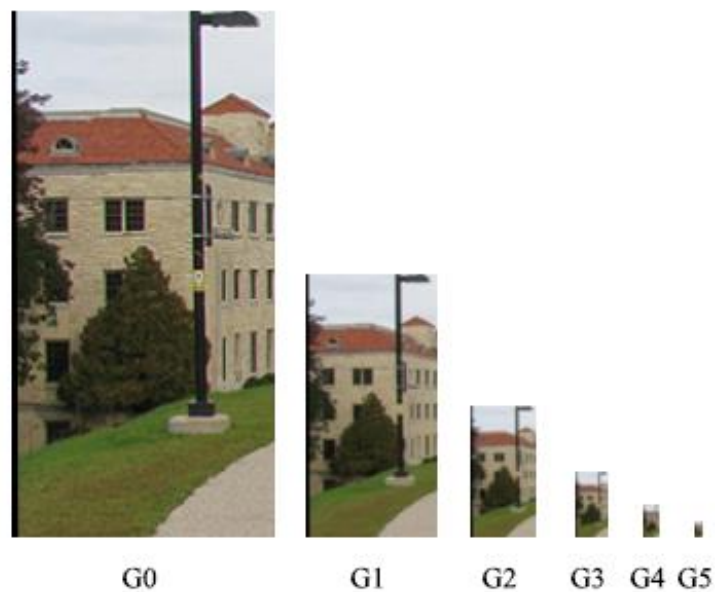


Fig. 14 Pirámide gaussiana derecha

El siguiente paso es calcular la pirámide laplaciana. El último nivel (en nuestro ejemplo el 5 se define como la imagen del último nivel de la pirámide gaussiana). Ahora vamos bajando de niveles de modo que cada nivel corresponde a la imagen del mismo nivel de la pirámide gaussiana restándole aquella que se obtiene como resultado de expandir el nivel actual+1 de la gaussiana. Como resultado se obtiene una imagen con las frecuencias altas. En las imágenes 15 y 16 se puede ver las pirámides laplacianas separadas de las dos imágenes de entrada.

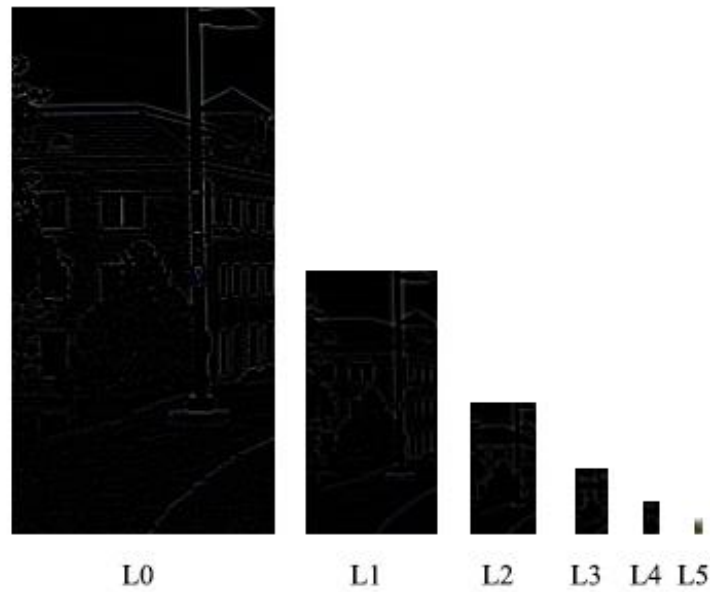


Fig. 15 Pirámide laplaciana izquierda

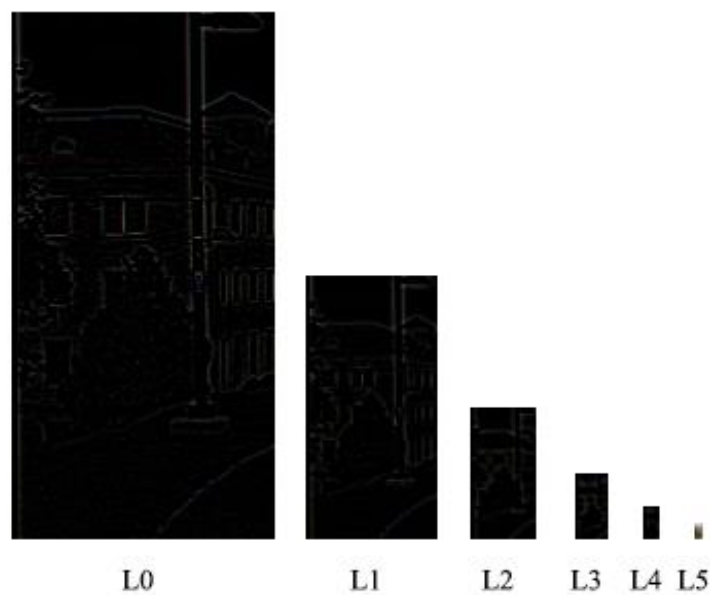


Fig. 16 Pirámide laplaciana derecha

A continuación simplemente creamos una nueva pirámide laplaciana cuyas imágenes están compuestas por la mitad izquierda de la imagen de la izquierda y la mitad derecha de la imagen de la derecha. Sería el equivalente a partir por la mitad las imágenes obtenidas en las dos pirámides laplacianas y juntar sus mitades [Fig 17].

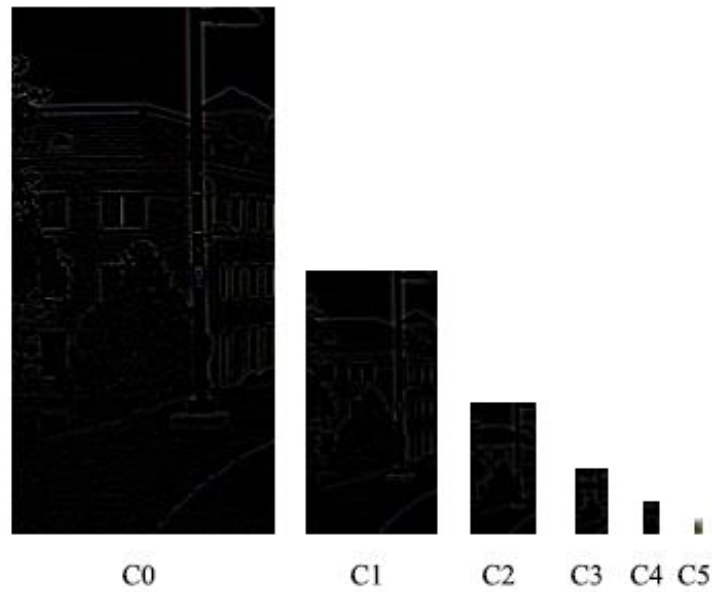


Fig. 17 Pirámides laplacianas combinadas

El último paso de la mezcla piramidal consiste en sumarle a la imagen de cada nivel de la pirámide laplaciana con las mitades combinadas la imagen resultante de expandir nivel+1 de dicha pirámide. En este caso nos servimos de los resultados parciales obtenidos de modo que actualizamos el valor del nivel actual de la laplaciana para considerarlo en la siguiente iteración. Para el último nivel de la pirámide (en nuestro caso el 5) no se realiza ningún cálculo si no que se toma tal cual para el proceso descrito en los siguientes niveles [Fig. 18].

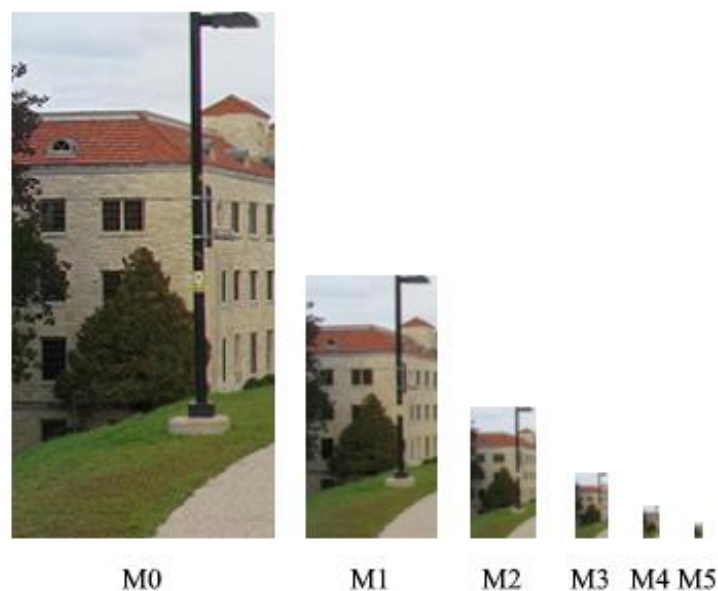


Fig. 18 Pirámide mezclada

Gracias a este método es posible una buena continuidad del color. En M0 tenemos la imagen mezclada obtenida de este ejemplo. Si nos fijamos en las imágenes de partida es apreciable que la de la izquierda tiene unos tonos mucho más fríos que la de la derecha. Sin embargo, como resultado de aplicar la mezcla piramidal tenemos una transición tan suave de los tonos fríos a los cálidos que es difícilmente apreciable.



## PROS Y CONTRAS

### PROS

Los mayores beneficios de esta técnica vienen del uso del algoritmo de mezcla usando el algoritmo de Burt-Adelson y del uso de la transformación a cilíndricas:

- Las transiciones entre imágenes son tremendamente suaves haciendo que sea realmente difícil apreciar el salto de una a otra.
- Funciona mucho mejor en imágenes con pequeñas variaciones de exposición que otras técnicas conocidas. [Fig. 19]
- Permite construir panoramas más largos que usando proyección de lineal sin apenas distorsión. [Fig. 20 y 21]



Fig. 19 Diferencia entre imagen sin algoritmo de mezcla (Izquierda) y con algoritmo de mezcla (derecha)



Fig. 20 Proyección perspectiva



Fig. 21 Proyección cilíndrica

## CONTRAS

Por el contrario, las limitaciones de esta técnica son:

- Estimar de forma automática las correspondencias con SIFT requiere gran capacidad de cómputo.
- Para la proyección cilíndrica se ha de usar un factor de escala que es desconocido. Y sin posibilidad de estimarlo de forma automática, para cada conjunto de imágenes, se deben probar distintos valores hasta encontrar el apropiado
- El paso a cilíndricas añade aliasing que se ve acentuado en imágenes con líneas rectas y con valores de escala bajos [Fig. 22].
- Las dos regiones que se unen por el algoritmo de mezcla han de ser lo suficientemente amplias como para que el algoritmo pueda mezclar las imágenes de forma eficaz.
- Si hay objetos en movimiento en las zonas de mezclado pueden aparecer artefactos como objetos borrosos o duplicados [Fig. 23].

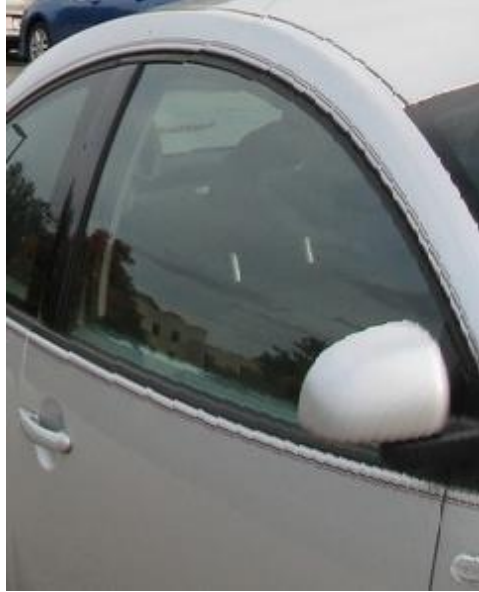


Fig. 22 Aliasing al aplicar la proyección cilíndrica



Fig. 23 Personas borrosas por estar en movimiento

#### POSIBLES PROPUESTAS DE MEJORA

Se proponen como mejoras a la técnica lo siguiente:

- Corregir el aliasing que aparece al aplicar la transformación cilíndrica usando alguna interpolación como la bilineal o la bicúbica.
- Estimación del factor de escala de la transformación cilíndrica de forma automática.
- Búsqueda automática de correspondencias y zonas a mezclar sin importar el orden de entrada de las imágenes, librando al usuario del trabajo de ordenar las imágenes.
- Corrección automática de la exposición y balance de blancos para que las imágenes no tengan cambios en la iluminación.
-

## RESULTADOS FINALES

A continuación se muestran un par de panoramas construidos con nuestro algoritmo:



## MANUAL DE USUARIO

El programa se ejecuta desde la línea de comandos para facilitar la selección de las imágenes e introducirlas en el programa de forma sencilla. Tiene tres parámetros que han de ser introducidos en el siguiente orden:

- Directorio de fotos.- Ruta hasta la carpeta con las fotos.
- Factor de escalado.- Valor para el factor de escala en la transformación a cilíndricas.
- Imagen de salida.- Parámetro opcional. Nombre de la imagen con la que se almacenará en disco el panorama

Algunas advertencias:

- Las fotos han de estar enumeradas en el orden exacto en el que se relacionan con la siguiente.
- Se recomienda encarecidamente usar el parámetro de “Imagen de salida”. Si no se usa, la imagen se mostrará en pantalla, pero debido a las limitaciones de OpenCV con el manejo de las ventanas, es muy probable que con panoramas grandes, la ventana corte la imagen.

Ejemplo de uso del programa:

```
Panorama.exe c:\imgs 900 panorama.jpg
```

## REFERENCIAS

<https://wiki.colby.edu/display/~hnphan/CS398+Final+Project+-+Cylindrical+Panorama>

[http://pages.cs.wisc.edu/~csverma/CS766\\_09/ImageMosaic/imagemosaic.html](http://pages.cs.wisc.edu/~csverma/CS766_09/ImageMosaic/imagemosaic.html)