



Ciencia y Tecnología

Secretaría de Ciencia, Humanidades, Tecnología e Innovación



CIMAT
UNIDAD MÉRIDA

SUBSTRACCIÓN DE FONDO PARA CÁMARAS EN MOVIMIENTO

Dr. Francisco J. Hernández López
SECIHTI – CIMAT-Mérida
fcoj23@cimat.mx, www.cimat.mx/~fcoj23



CLASIFICACIÓN DE VIDEOS

- Es posible caracterizar el dominio de los videos en 4 clases, dependiendo del movimiento de la cámara y el tipo de procesamiento

	Fuera de Línea (offline)	En Línea (online)
Cámara Estática	Pos-procesamiento	Video vigilancia
Cámara Móvil	Youtube	Celulares, TV, Robótica, Autos, Google Glass, ...

MÉTODOS CONSIDERANDO CÁMARAS EN MOVIMIENTO

- Los algoritmos de substracción de BG tradicionales, no son aplicables directamente al caso de cámaras en movimiento
- Existen algunas extensiones de los algoritmos tradicionales, que consisten en utilizar compensación de movimiento
- Además se han introducido muchos algoritmos para detección de FG de objetos específicos, usando datos de entrenamiento, es posible tener detectores especializados:

- Peatones
- Rostros
- Vehículos
- Bicicletas
- Etc.

Caltech Pedestrian Detection Benchmark



CLASIFICADOR DE OBJETOS

- Dado un clasificador entrenado para un cierto objeto
- Para encontrar las ocurrencias del objeto en la imagen, la idea es escanear la imagen y aplicar el detector en cada posición y escala



Deformable Part Model (DPM) with motion feature on
Caltech Pedestrian Dataset 1

ALGORITMOS DE DETECCIÓN DE OBJETOS ESPECÍFICOS

- Para el caso de videos a partir de cámaras fijas, estos algoritmos están lejos de lograr el rendimiento que tienen los algoritmos típicos de substracción de fondo, por ejemplo en términos de las curvas ROC
- No proveen una segmentación exacta del objeto, pero si dan una caja envolvente (bounding box) alrededor del objeto
- Para el caso de videos a partir de cámaras móviles, estos algoritmos se aplican directamente, no asumen cámaras o escenas estacionarias y se aplican a cada frame del video de forma independiente

ALGORITMOS DE DETECCIÓN DE OBJETOS ESPECÍFICOS (C1)

- Limitaciones que los hacen no efectivos:

1. Diseñados para detectar objetos a partir de imágenes, no usan restricciones temporales en el video
2. Aún para imágenes, el rendimiento de estos algoritmos nos dice que están lejos de llegar a ser utilizados en aplicaciones del mundo real: 47.3% de precisión

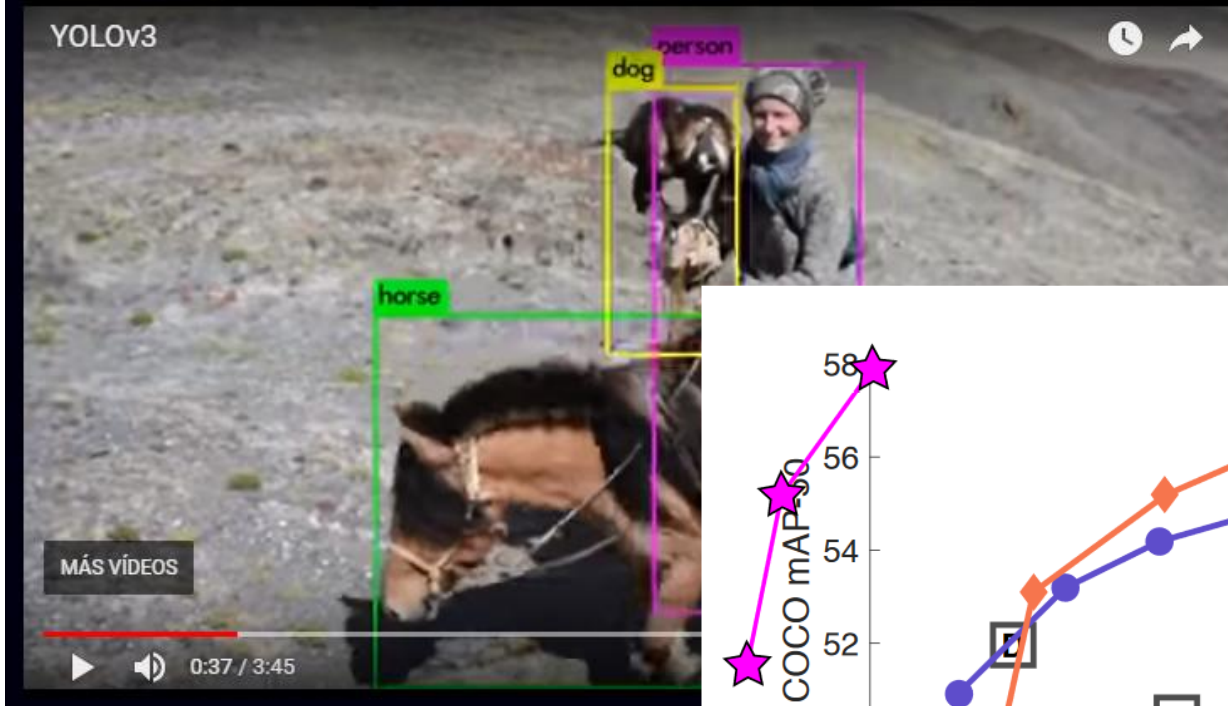
		back	aero											dining	motor		potted			tv/		
	[mean]	ground	plane	bicycle	bird	boat	bottle	bus	car	cat	chair	cow	table	dog	horse	bike	person	plant	sheep	sofa	train	monitor
BONNGC O2P CPMC CSI	45.4	85.0	59.3	27.9	43.9	39.8	41.4	52.2	61.5	56.4	13.6	44.5	26.1	42.8	51.7	57.9	51.3	29.8	45.7	28.8	49.9	43.3
BONN CMBR O2P CPMC LIN	44.8	83.9	60.0	27.3	46.4	40.0	41.7	57.6	59.0	50.4	10.0	41.6	22.3	43.0	51.7	56.8	50.1	33.7	43.7	29.5	47.5	44.7
BONN O2PCPMC FGT SEGM	47.0	85.1	65.4	29.3	51.3	33.4	44.2	59.8	60.3	52.5	13.6	53.6	32.6	40.3	57.6	57.3	49.0	33.5	53.5	29.2	47.6	37.6
NUS DET SPR GC SP	47.3	82.8	52.9	31.0	39.8	44.5	58.9	60.8	52.5	49.0	22.6	38.1	27.5	47.4	52.4	46.8	51.9	35.7	55.3	40.8	54.2	47.8
UVA OPT NBNN CRF	11.3	63.2	10.5	2.3	3.0	3.0	1.0	30.2	14.9	15.0	0.2	6.1	2.3	5.1	12.1	15.3	23.4	0.5	8.9	3.5	10.7	5.3

<http://host.robots.ox.ac.uk:8080/pascal/VOC/voc2012/results/index.html>

3. Escalar estos algoritmos para un número grande de clases y alcanzar una buena precisión aún es un reto

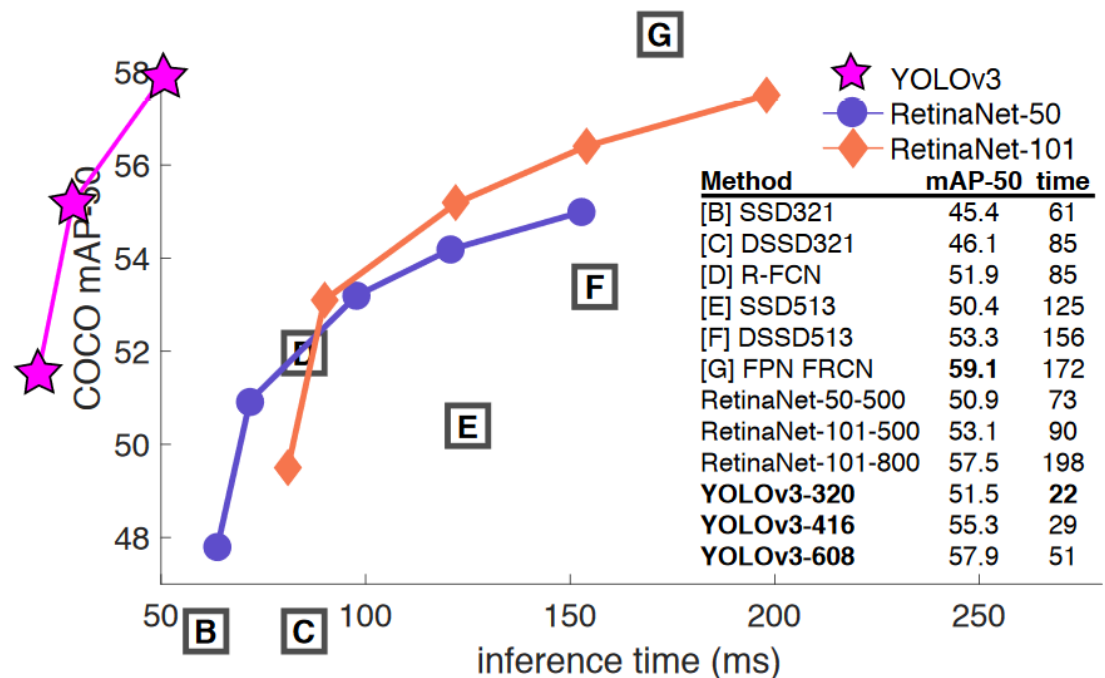
YOLO: Real-Time Object Detection

You only look once (YOLO) is a state-of-the-art, real-time object detection system. On a Pascal Titan X it processes images at 30 FPS and has a mAP of 57.9% on COCO test-dev.



<https://pjreddie.com/darknet/yolo/>

Dataset: Microsoft COCO



¿QUÉ ES FG? Y ¿QUÉ ES BG?

- En el caso de cámaras moviéndose libremente, la definición de FG y BG puede ser ambigua
- Un método de substracción de fondo para cámaras móviles debería usar e integrar varias señales perceptuales para segmentar el FG, entre ellas:
 - **Discontinuidad de movimiento** → Regiones de la imagen y objetos que se mueven de forma independiente del resto de la escena
 - **Discontinuidad de profundidad** → Objetos que están más cerca de la cámara que el resto de la escena
 - **Discontinuidad de apariencia** → Objetos salientes o regiones de la imagen que se observan diferente a partir del entorno que lo rodea
 - **Formas familiares** → Regiones de la imagen que lucen como ciertas formas similares: personas, vehículos, animales, etc.

TÉCNICAS DE SUBTRACCIÓN DE BG USANDO COMPENSACIÓN DE MOVIMIENTO

- Estiman el movimiento o la transformación geométrica entre frames consecutivos o entre cada frame y un mapa del BG extendido ([Mosaico o Panorama](#))
- Típicamente usan una Homografía o una transformación afín 2D
- La exactitud de la estimación del movimiento depende:
 - El modelo de movimiento vaya de acuerdo con la escena observada
 - De la existencia de objetos del FG, lo cuales no siguen la suposición del modelo de movimiento
- Si el movimiento de la cámara es una rotación pura o solo zoom (sin traslaciones), entonces la transformación geométrica entre las imágenes puede ser modelada con una Homografía sin necesidad de suponer que la escena es plana

PANORAMA

- Panorama se deriva de las palabras griegas:
 - pân (= todo)
 - hòrama (= ver)
- Consiste en pegar los frames de un video (o las fotos de una escena) en una sola imagen, aumentando el campo de vista de la cámara



HISTORIA

- **1787**, Robert Barker (Pintor Irlandés) patentó sus ideas para una construcción panorámica cilíndrica de sus pinturas.
- **1793**, Robert Barker inauguró el primer edificio circular en Londres, llamada “Rotunda”.
 - Desde entonces se mostraban cada año en Leicester Square (centro del cine en Londres, Inglaterra) uno o dos nuevos panoramas.
 - Los panoramas eran pintados en lienzos (llegaron a medir aprox. 15m de altura x 100m de largo)



London desde el techo del edificio Albion Mills

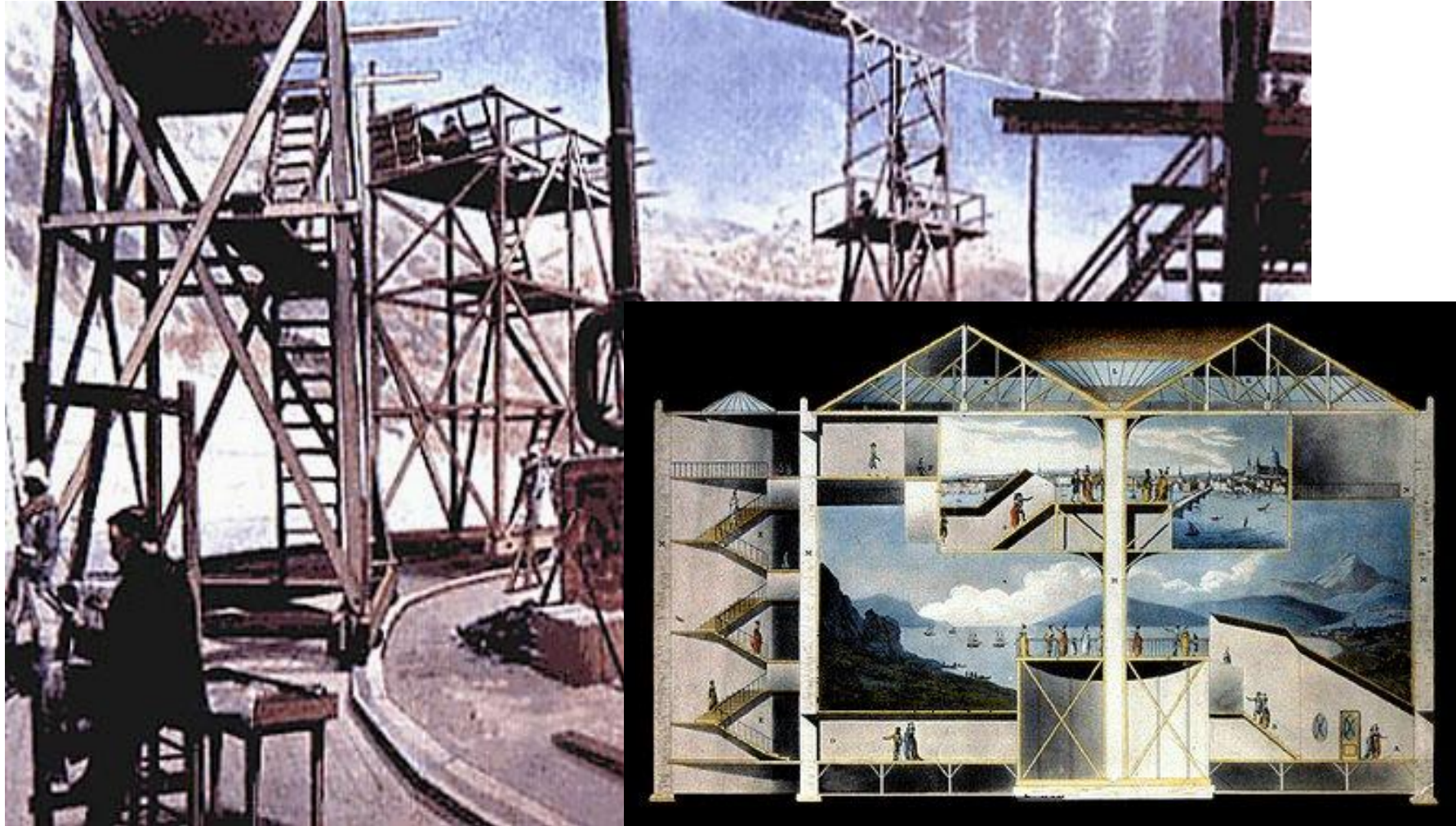
A History of Panoramic Image Creation, Harald Woeste

<http://www.graphics.com/article-old/history-panoramic-image-creation>

Substracción de fondo, cámaras en movimiento. Francisco J. Hernández-López

Ene-Jun 2025

ROTUNDA



27m de diámetro, 17m de altura

La visión total: el Panorama de Barker, Publicado el 13/06/2011 por Savinio

Substracción de fondo, cámaras en movimiento. Francisco J. Hernández-López

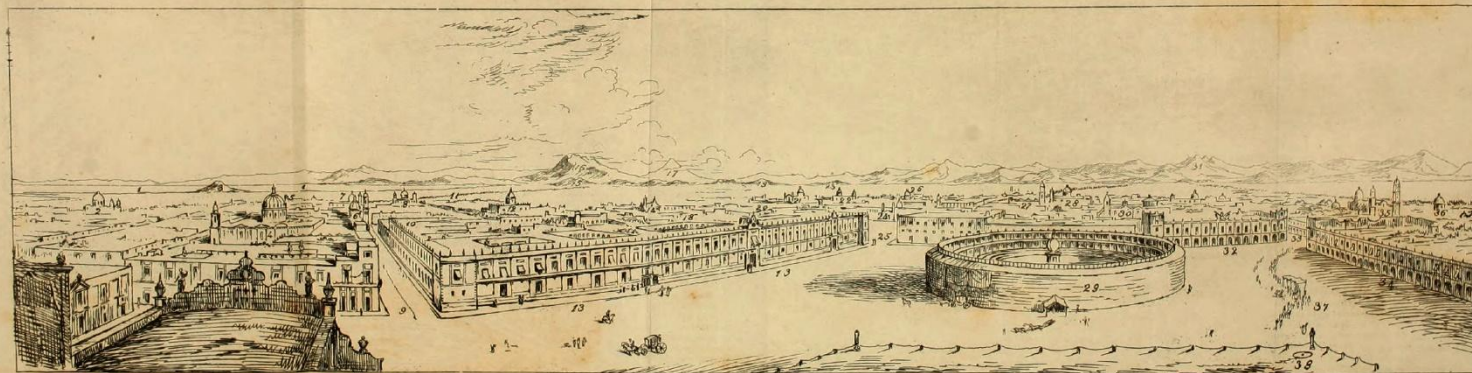
Ene-Jun 2025

PANORAMA DE LA CIUDAD DE MÉXICO

EXPLANATION of a VIEW of the CITY of MEXICO, exhibiting in the PANORAMA, LEICESTER SQUARE.



- | | | | | | |
|-------------------------------------|--|-----------------------------|--|--|---|
| 1. Calle Potosí. | 7. Colegio de las Niñas Hospital for poor Girls. | 13. S ^a Ypolito. | 19. San Lorenzo. | 24. R ^o Alamo, Custom House. | 30. La Encarnacion. |
| 2. De los Virreyes. | 8. S. Francisco. | 14. S ^a Clara. | 20. Calle de Tacuba. | 25. Casa de l'Inquisicion. | 31. Mountain of Chiquita. |
| 3. Calzaday caneria de Chapultepec. | 9. Le Profeso. | 15. S. Fernando. | 21. Imprinta, general Printing Office. | 26. S ^a Catalina de Sena. | 32. Calzaday miera de Guadalupe. |
| 4. Chapultepec. | 10. La Alameda. | 16. Calle Lerrea. | 22. Cruz taba verterus. | 27. Plaza de Galeis. | 33. Piramide de S ^a Juan. |
| 5. La Cagurros. | 11. Mineria. | 17. Casa de Stada. | 23. First Chapel built by Cortez. | 28. Claverio, Treasury of the Cathedral. | 34. S. Pedro y S. Pablo in which the Congress meet. |
| 6. Espiritu Santo. | 12. Belemitas Convent for Friars. | 18. Concepcion. | 25. S. Domingo. | 29. Cathedral. | 35. S. Valdezeno. |



- | | | | | | |
|---|-----------------------------------|----------------------------|------------------------|--------------------------------------|-------------------------------------|
| 1. Lake of Texcoco. | 7. La S. S ^a Trinidad. | 13. Palacio. | 19. Lieta de Molango. | 25. Del Bolador Market. | 31. Lieta de Apasco. |
| 2. Nuestra Señora de Loreto. | 8. Soledad de Santa Cruz. | 14. La Merced. | 20. Universidad. | 26. S. Jaro de Gracia. | 32. Deposition, House of the Mayor. |
| 3. Peñon de los Banos. | 9. Calle del Arzobispado. | 15. Penon Vieque. | 21. S. Pablo. | 27. Hosp. de Jesus de los naturales. | 33. Calle Monterilla. |
| 4. S ^a Teresa de Antigua. | 10. Casa de Moneda. | 16. Islacahuatl. | 22. Col. de S. Pablo. | 28. S. Miguel. | 34. Los Portales. |
| 5. S. Lazaro. | 11. Calzaday. | 17. Popocatepetl. | 23. Estapalapa. | 29. Casa de Toros. | 35. S. Augustin. |
| 6. S ^a Ther, Convent for Nuns. | 12. Jesus Maria. | 18. Botanic Garden. | 24. Porta Gali. | 30. S. Bernardo. | 36. Regina Celi. |
| | | 37. Religious Processions. | 38. Sacrificial Stone. | | |

En 1825 se mostraron los panoramas de Mexico en la Rotunda.

<http://www.redalyc.org/articulo.oa?id=60021670002>

Substracción de fondo, cámaras en movimiento. Francisco J. Hernández-López

Ene-Jun 2025

DAGUERROTIPO

- **1833** Daguerre da a conocer oficialmente su procedimiento para obtener una fotografía a partir de una superficie de plata pulida (“Daguerrotipo”)



<http://es.wikipedia.org/wiki/Daguerrotipo>

- **1851** Martin Behrmanx crea una vista panorámica de San Francisco tomada desde Rincon Hill usando el Daguerrotipo. El panorama fue creado con 5 placas Daguerrotipos.



http://en.wikipedia.org/wiki/Panoramic_photography

- El tiempo de exposición o toma de la fotografía era en minutos.

COLODIÓN HÚMEDO

- **1851**, Gustave Le Gray. Se habla por primera vez de fotografía “instantánea”.
- Ahora el tiempo de exposición podía ser de un poco menos de 30 segundos para obtener una fotografía.



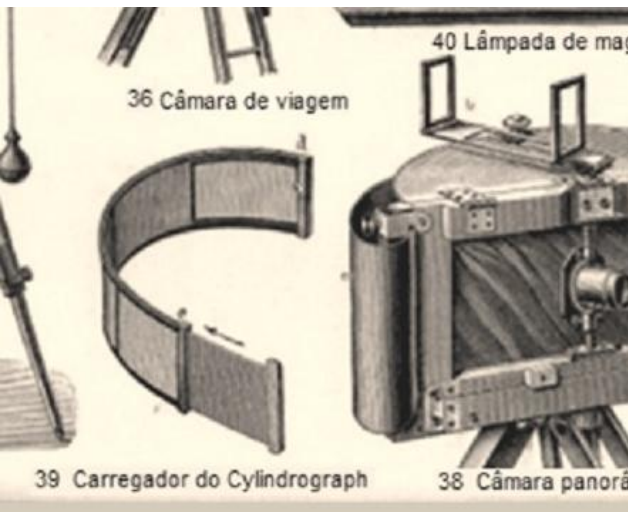
Lookout Mountain, Tennessee, **1864**, George N. Barnard



Caballo sin cabeza, Woodstock, US, **1870**

PRIMERAS CÁMARAS PANORÁMICAS

- **1888**, aparecen en el mercado:
 - Cylindrograph, Wonder Panoramic, Pantascopic, Cyclo-Pan, Kodak etc.



Cylindrograph



Pantascopic

Kodak: "You press the button, we do the rest"

CÁMARAS DIGITALES

- **1975**, Steve Seasson, ingeniero de Kodak crea la cámara digital.



- **1994-1996**, Aparecen las cámaras que podías conectar a la PC: Apple QuickTake 100, Kodak DC40, Casio QV-11, Sony's Cyber-Shot Digital.



- **2000**, aparece el primer teléfono celular comercial con cámara J-SH04, hecho por Sharp Corporation, Japón.

MÉTODOS PARA OBTENER UN PANORAMA

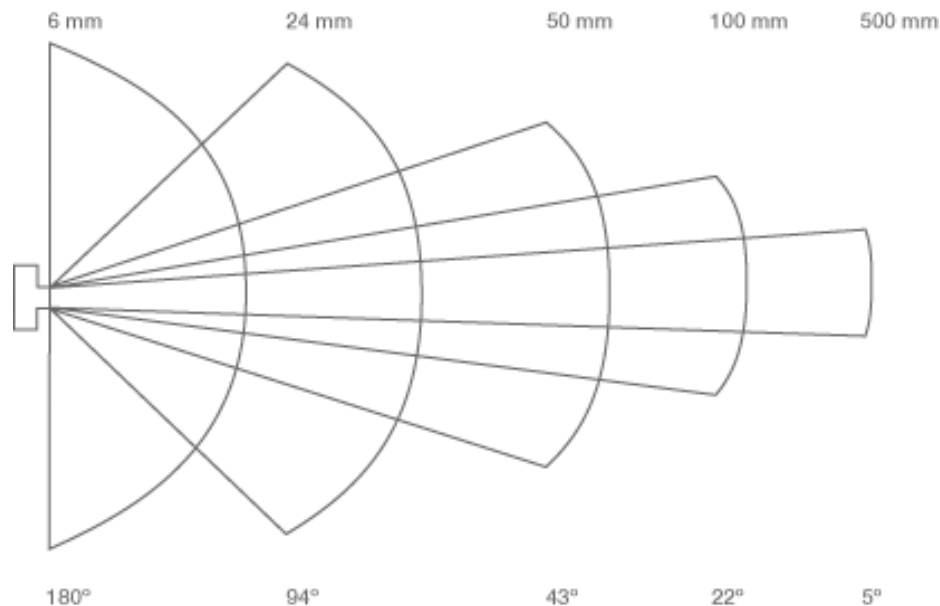
- **Rotación corta:** Hay cámaras que tienen una lente que gira alrededor del punto nodal trasero de la cámara y proyecta la imagen sobre una película curva. Campo de vista: 110°-140°.
- **Rotación completa:** Un mecanismo gira la cámara de forma continua y va jalando la película a través de la cámara, de tal forma que el movimiento de la película coincide con el movimiento de la imagen. Campo de vista: 360°.



Panoramic photography,
http://en.wikipedia.org/wiki/Panoramic_photography

MÉTODOS PARA OBTENER UN PANORAMA (C1)

- **Lente fijo o visión amplia:** El campo de vista más amplio que puede ser capturado con enfoque y con poca distorsión y viñeteado es de 90° . $> 120^\circ$ se requiere de un filtro especial para corregir el viñeteado en los bordes. Lentes con ángulos de hasta 180° (ojo de pescado), presentan extrema distorsión geométrica.



Panoramic cameras,
<https://ppyrx.wordpress.com/2015/01/09/panoramic-cameras/>

MÉTODOS PARA OBTENER UN PANORAMA (C2)

- **Cámaras catadiópticas:** Consisten en una lente y un espejo



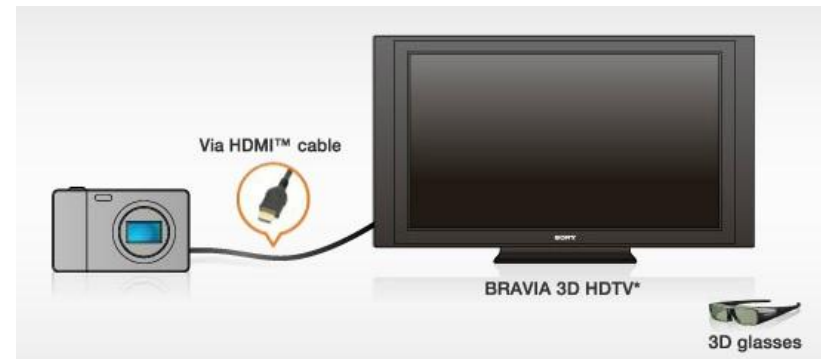
http://www1.cs.columbia.edu/CAVE/projects/cat_cam_360/

- **Panorama 3D:** Tomando el video desde la cámara con dos lentes (par stereo) o con una lente y tomando dos frames (a la vez) a diferente enfoque.



Panoramic photography,

http://en.wikipedia.org/wiki/Panoramic_photography



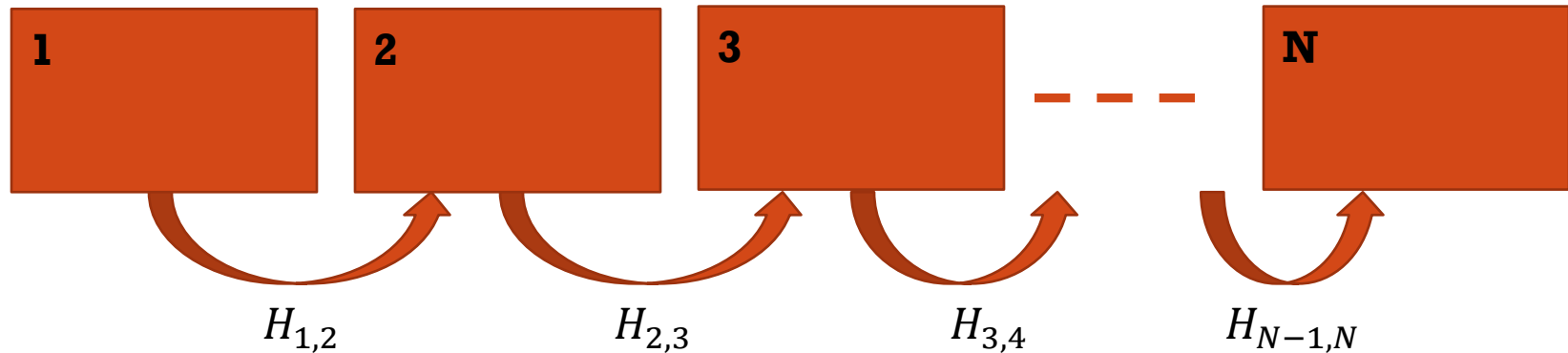
MÉTODOS PARA OBTENER UN PANORAMA (C3)

- **Foto stitching:** La idea es tomar una serie de imágenes y unir las mediante algún software computacional.



CONSTRUCCIÓN DE PANORAMAS

CALCULAR LA HOMOGRAFÍA ENTRE CADA PAR SUBSECUENTE DE IMÁGENES DE UN VIDEO



- Homografía de la imagen t con respecto a la imagen de referencia 1:

$$H_{1,t} = H_{t-1,t} H_{1,t-1}, \text{ con } H_{1,1} = I \text{ y } t > 1$$

- Para transformar los puntos de una imagen a otra:

$$\Gamma(I, H, x) \stackrel{\text{def}}{=} B\{I, C(HG(x))\},$$

con

$G \rightarrow$ transforma un punto (pixel) a coordenadas homogneas

$C \rightarrow$ transforma coordenadas homogneas a coordenadas cartesianas

$B \rightarrow$ interpola los valores de I

COMPOSICIÓN DE IMÁGENES (C1)

- Elegimos una imagen de referencia I_r

- Cada imagen es trasformada con la homografía $H_{r,t}$

$$\hat{I}_t(x) = \begin{cases} \Gamma(I_t, H_{r,t}, x) & \text{si } H_{r,t}\hat{x} \in \mathcal{L} \\ \text{NAN} & \text{otro caso} \end{cases}$$

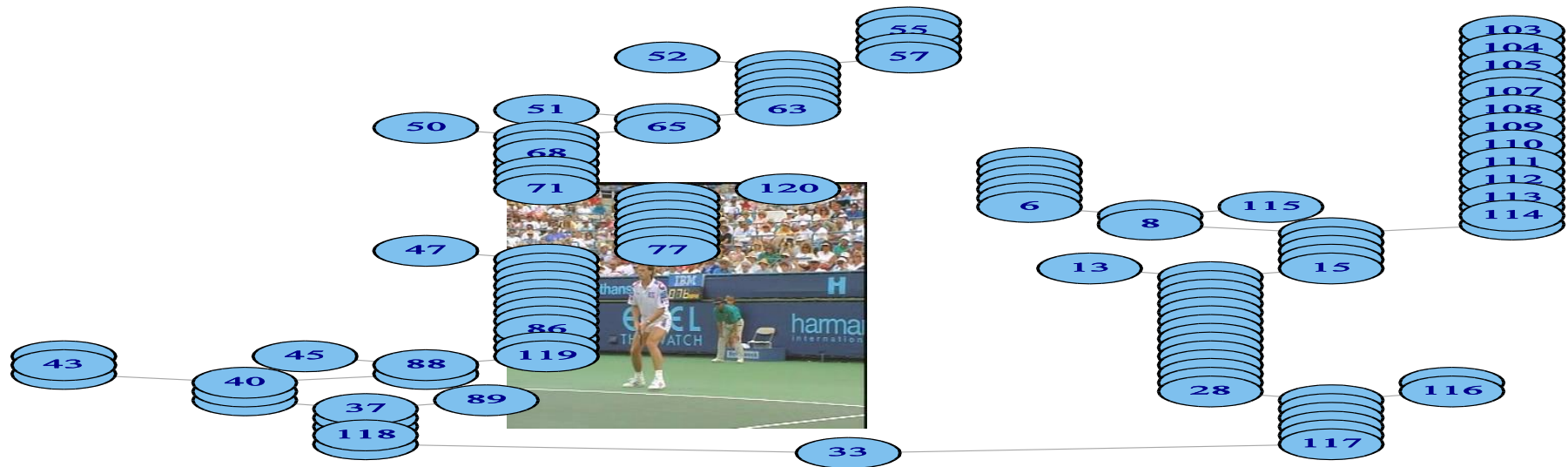
con $\mathcal{L}$ la malla de la imagen I_t

- El panorama o mosaico es generado con

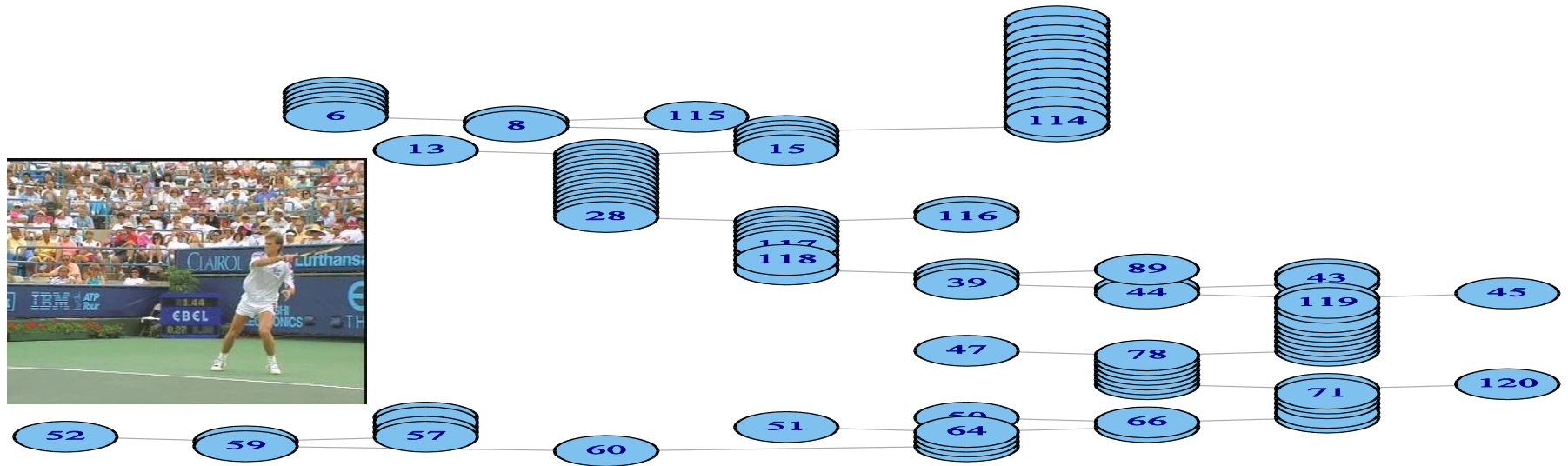
$$m_i(x) = \text{mediana}[\hat{f}_i^t(x)] \quad \forall i \in \{1,2,3\}, t = 1, \dots, N$$



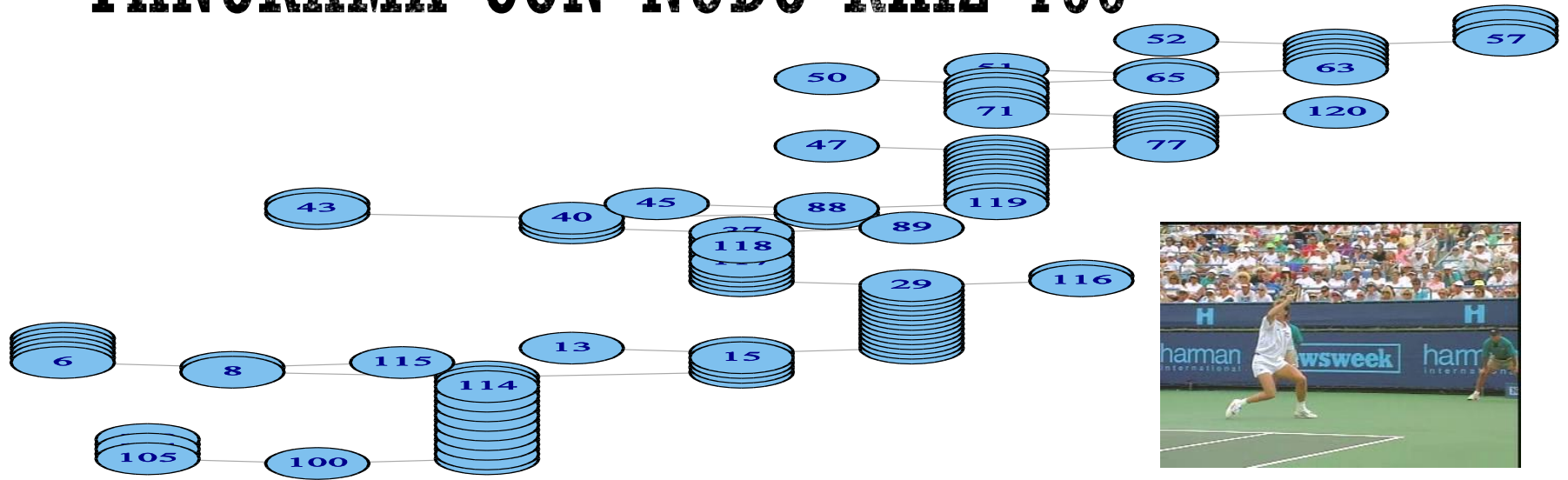
PANORAMA CON NODO RAÍZ 33



PANORAMA CON NODO RAÍZ 60



PANORAMA CON NODO RAÍZ 100



OPENCV STITCHING, --WARP PLANE



OPENCV STITCHING, --WARP CYLINDRICAL



OPENCV STITCHING, --WARP SPHERICAL



OPENCV STITCHING, --WARP FISHEYE



OPENCV STITCHING, --WARP PANINI PORTRAITS 1.5B1



OPENCV STITCHING, --WARP MERCATOR



TRABAJO RELACIONADO

- Bartoli et al. 2002



Adrien Bartoli, Navneet Dalal, Biswajit Bose and Radu Horaud. From Video Sequences to Motion Panoramas. In Proceedings of the Workshop on Motion and Video Computing, MOTION '02, pages 201. IEEE Computer Society, 2002

TRABAJO RELACIONADO (C1)

- Farin et al. 2004



(a) sprite 1 (frames 1–242)

Solo hasta 180° del campo de vista puede ser representado en un panorama (sprite)

Modelo de movimiento:
Homografía H



(b) sprite 2 (frames 243–255)

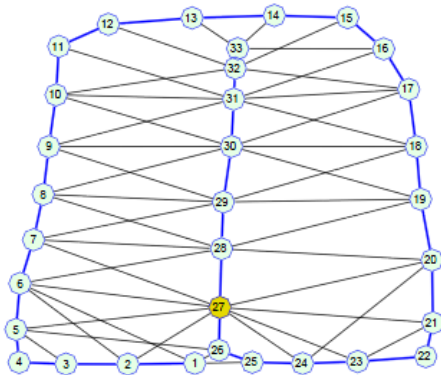


(c) sprite 3 (frames 256–293)

Dirk Farin, WA Eelsberger et al. Video-object segmentation using multi-sprite background subtraction. In Multimedia and Expo, 2004. ICME'04. 2004 IEEE International Conference on, volume 1, pages 343346. IEEE, 2004

TRABAJO RELACIONADO (C2)

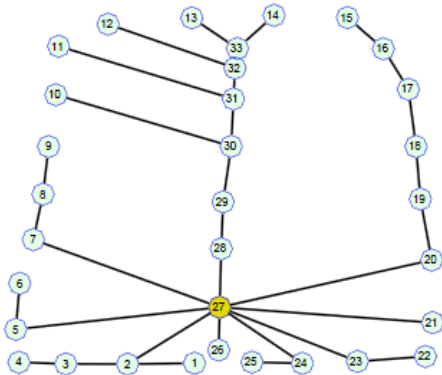
■ Marzotto et al. 2004



Supuestos:

- La escena es plana
- El punto de vista no cambia

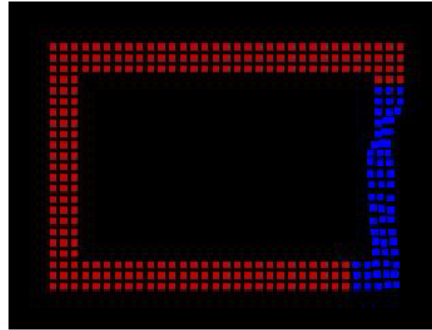
Modelo de movimiento:
Homografía H



Roberto Marzotto, Andrea Fusiello and Vittorio Murino. High resolution video mosaicing with global alignment. In Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on, volume 1, pages I692. IEEE, 2004

TRABAJO RELACIONADO (C3)

- Azzari et al. 2005, Bevilacqua et al. 2006



- Proponen usar un conjunto de puntos cerca de los bordes del frame como puntos característicos
- Usan clustering para separar el movimiento de la cámara del movimiento de los objetos



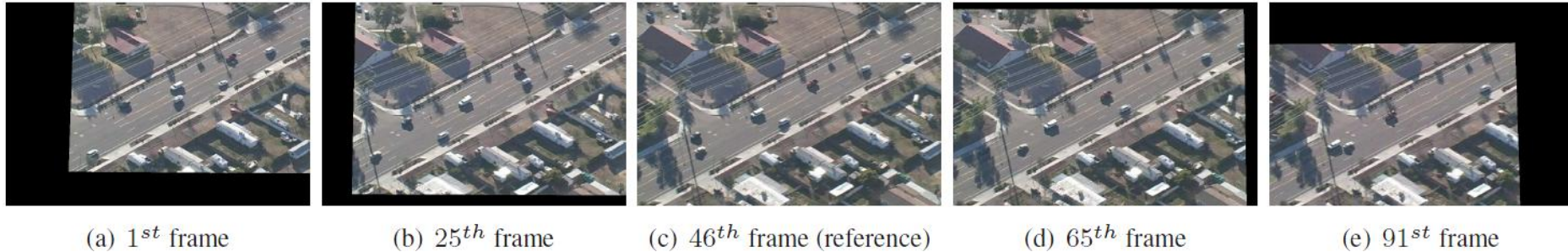
Modelo de
movimiento:
Homografía H

Procesan frames
de 320×240
píxeles a 11 fps

Pietro Azzari, Luigi di Stefano and Alessandro Bevilacqua. An effective real-time mosaicing algorithm apt to detect motion through background subtraction using a PTZ camera. In AVSS, pages 511516. IEEE Computer Society, 2005

TRABAJO RELACIONADO (C4)

- Yu et al. 2008



Modelo de movimiento:
Homografía H

Su implementación en GPU procesa videos de 320×240 pixels a 10 fps calculando la moda de 91 frames



Qian Yu and Gérard Medioni. A GPU-based implementation of Motion Detection from a Moving Platform. In Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference on, pages 16. IEEE, 2008

TRABAJO RELACIONADO (C5)

- Colombari et al. 2007

Con base en el trabajo de Marzotto et al. 2004 para calcular el panorama (usando grafos)



Andrea Colombari, Andrea Fusiello and Vittorio Murino. Segmentation and tracking of multiple video objects. Pattern Recogn., vol. 40, no. 4, pages 13071317, 2007

APLICACIONES

ROTUNDA, PINTURA CILÍNDRICA



Panorama Racławice (1893) localizado en Wrocław, Polonia. Mide aprox. 15m x 120m.

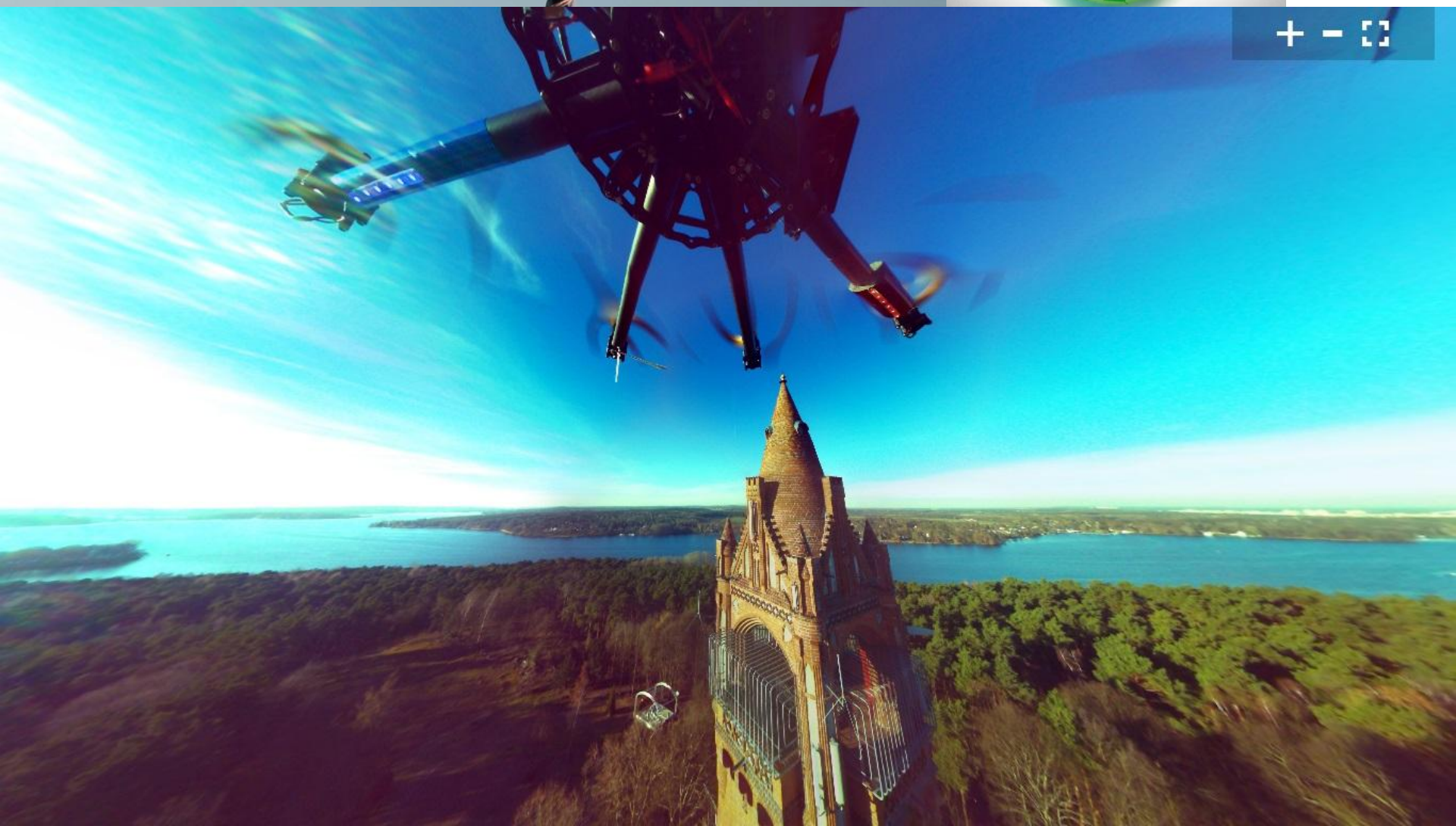
REALIDAD AUMENTADA



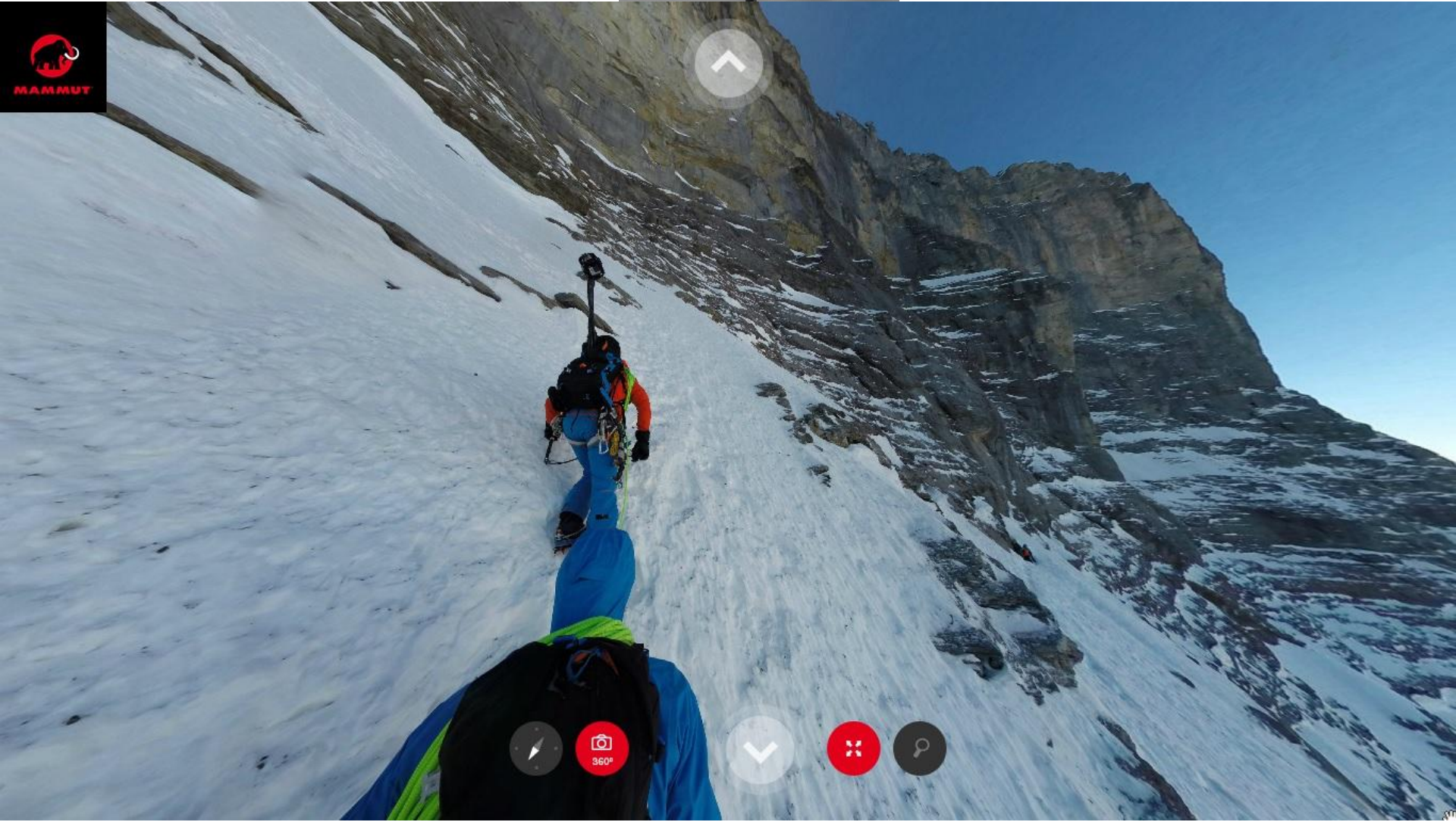
En el 2005, artistas exhibieron una pintura panorámica de 180°, mostrando la famosa puerta de Brandenburg a finales de la segunda guerra mundial.



PANONO
Panoramic Ball Camera

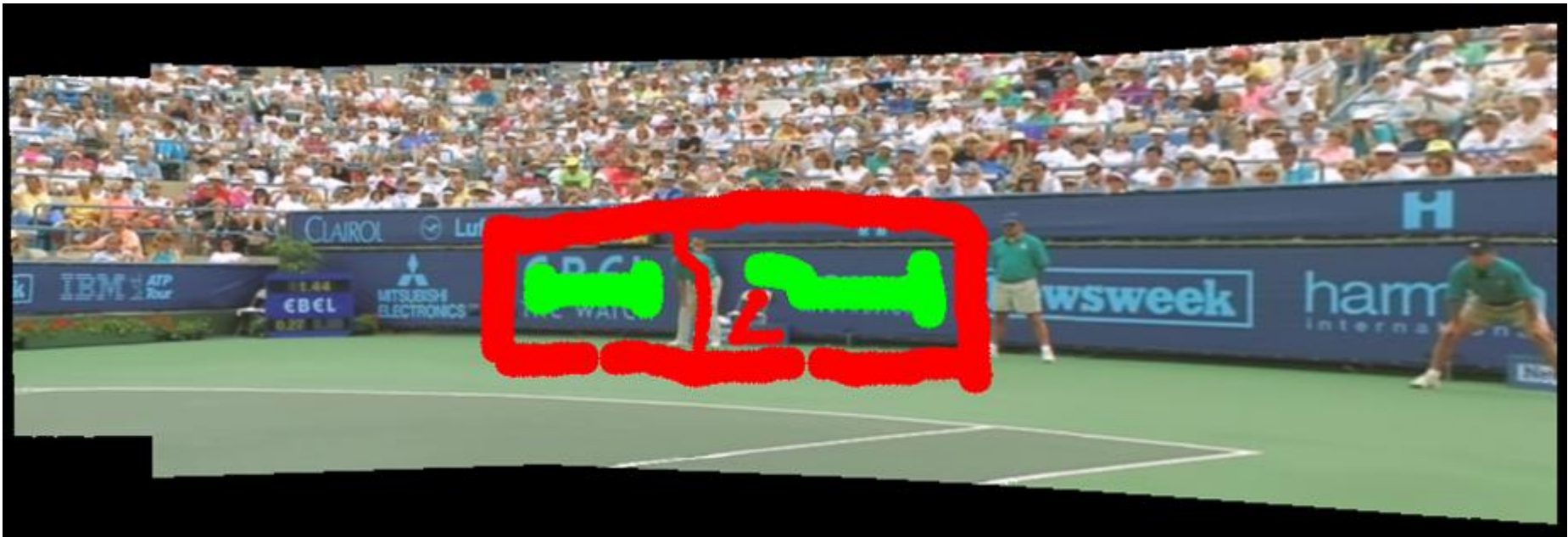


GOPRO 360



AVSCREEN USANDO EL PANORAMA

- Interacción con el usuario



[Hernandez] Francisco J. Hernandez-Lopez and M. Rivera. AVScreen: a Real-Time video augmentation method. J. of Real-Time Image Process., pages 1-13, 2013.

AVSCREEN USANDO EL PANORAMA (C1)

- Proceso automático



Video original



Video editado

AVSCREEN USANDO EL PANORAMA (C2)



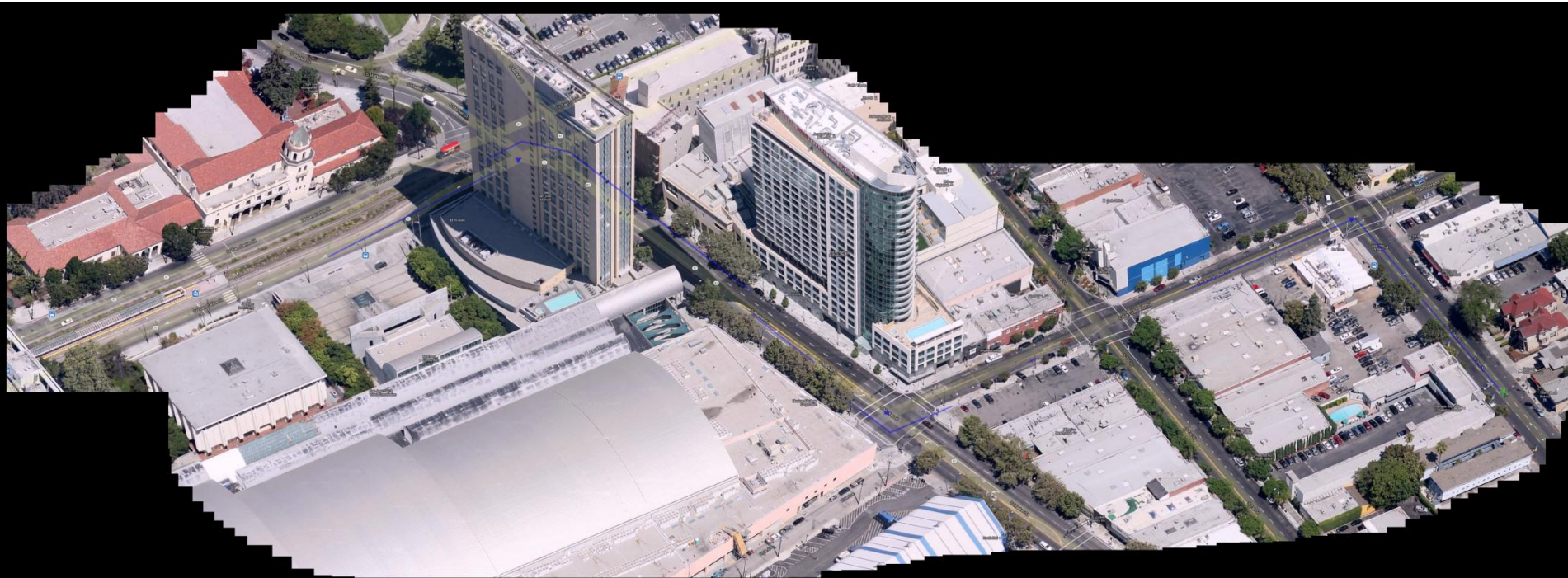
AVSCREEN USANDO EL PANORAMA (C3)



PANORAMA A PARTIR DE TOMAS AÉREAS



PANORAMA A PARTIR DE TOMAS AÉREAS (C1)



PANORAMA EN UN AMBIENTE 3D



CSI NY Temporada 4, Episodio 18: Admisiones

Substracción de fondo, cámaras en movimiento. Francisco J. Hernández-López

Ene-Jun 2025

GRACIAS POR SU ATENCIÓN

Francisco J. Hernandez-Lopez

fcoj23@cimat.mx

WebPage:

www.cimat.mx/~fcoj23

