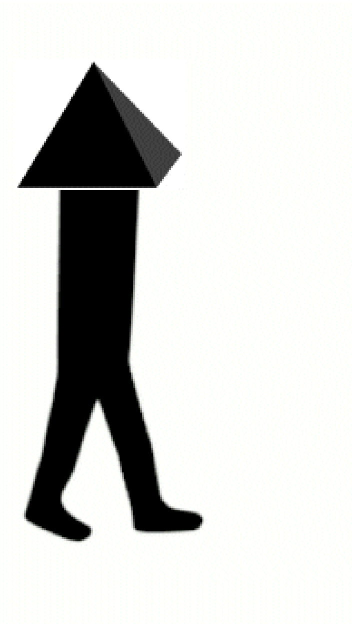
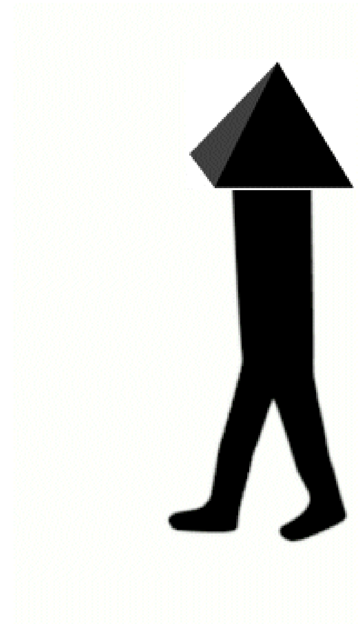
Representing Scenes with dynamic objects by Graph Pyramids

*When Pyramids Learned Walking (See [Kro09])**



Walter G. Kropatsch
PRIP, Vienna University of Technology, Austria

February 8, 2010



Contributions of Yll Haxhimusa, Adrian Ion, Nicole Artner, Helena Molina Abril, Mabel Iglesias-Ham, Dan Shao, Esther Antunez-Ortiz, Luis Alfredo Mateos are cordially acknwledged!

next: Contents

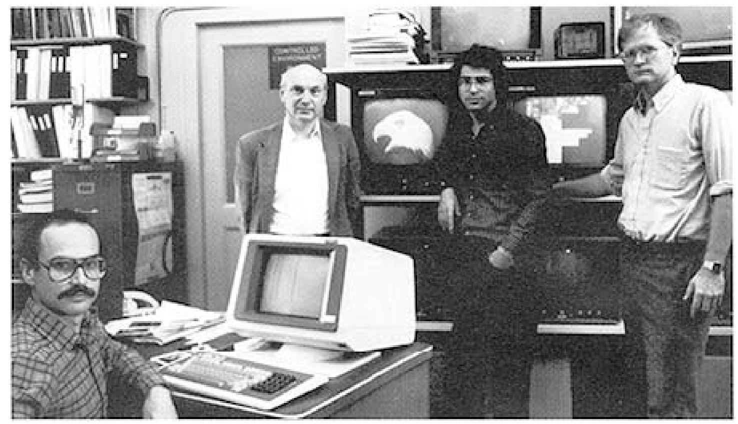
*Supported by the Austrian Science Fund under grants S9103-N13, P20134-N13 and P18716-N13.

Contents

- Pyramids
  - Regular Laplacian Pyramid
  - Structure Preserving Graph Pyramid
- Structure and Tracking
  - Single Target
  - Multiple Target Constellation
  - Tracking a Constellation with a Spring System
  - Motion Segmentation from Moving Triangulations
- The Walking Pyramid
- Conclusion and Outlook

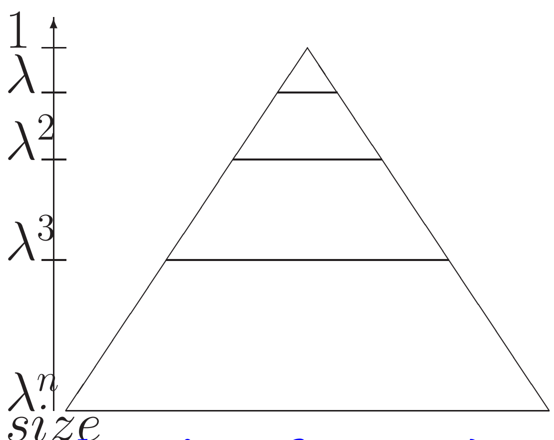
1984: Pyramid methods in image processing

- Bergen, Anderson, Adelson, Burt [AAB+84]
- **Gaussian** pyramid
- **(RE) Laplacian** pyramid
- low and band pass
- **image compression** 8:1 [BA83a]

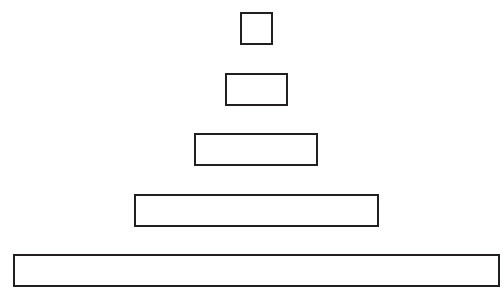


- multi-resolution **coring** reduces random noise and sharpens details
- multi-resolution **spline** fuses 'apple and orange' [BA83b]
- **change detection and tracking**[ABv85]
- pipelined pyramid machine

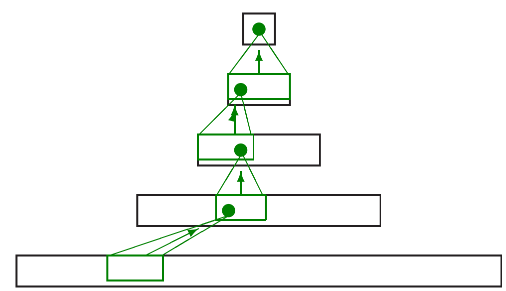
REGULAR Image PYRAMID



reduction factor $\lambda > 1.0$



Discrete levels



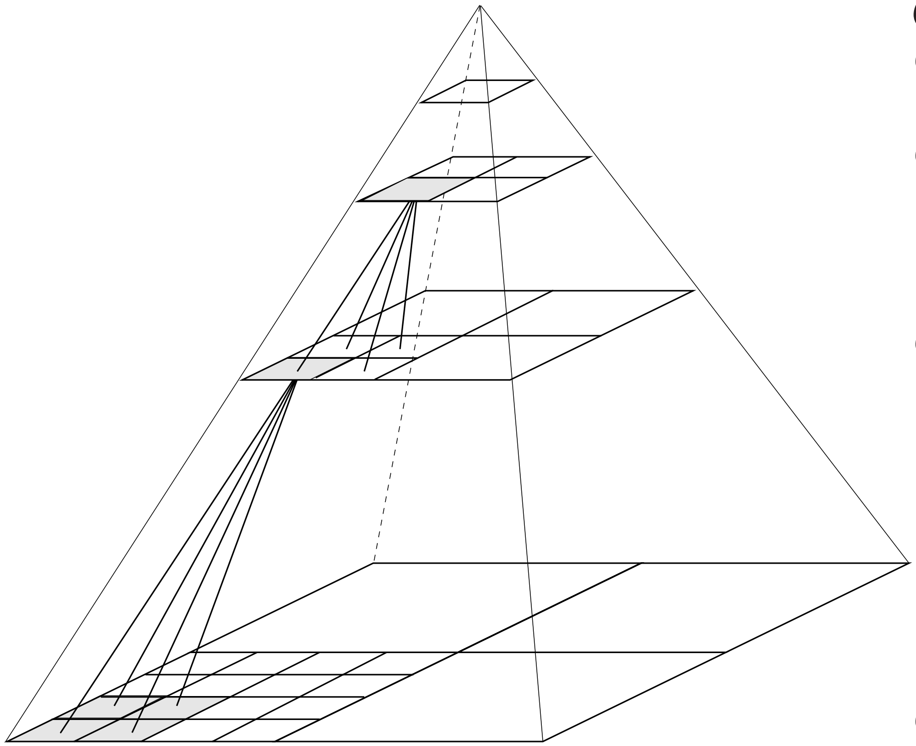
reduction window
reduction function

Major advantages:

$$\textbf{HEIGHT} \leq \frac{\log(image\_size)}{\log(\lambda)} = n.$$

Access to λ^n data in n steps.

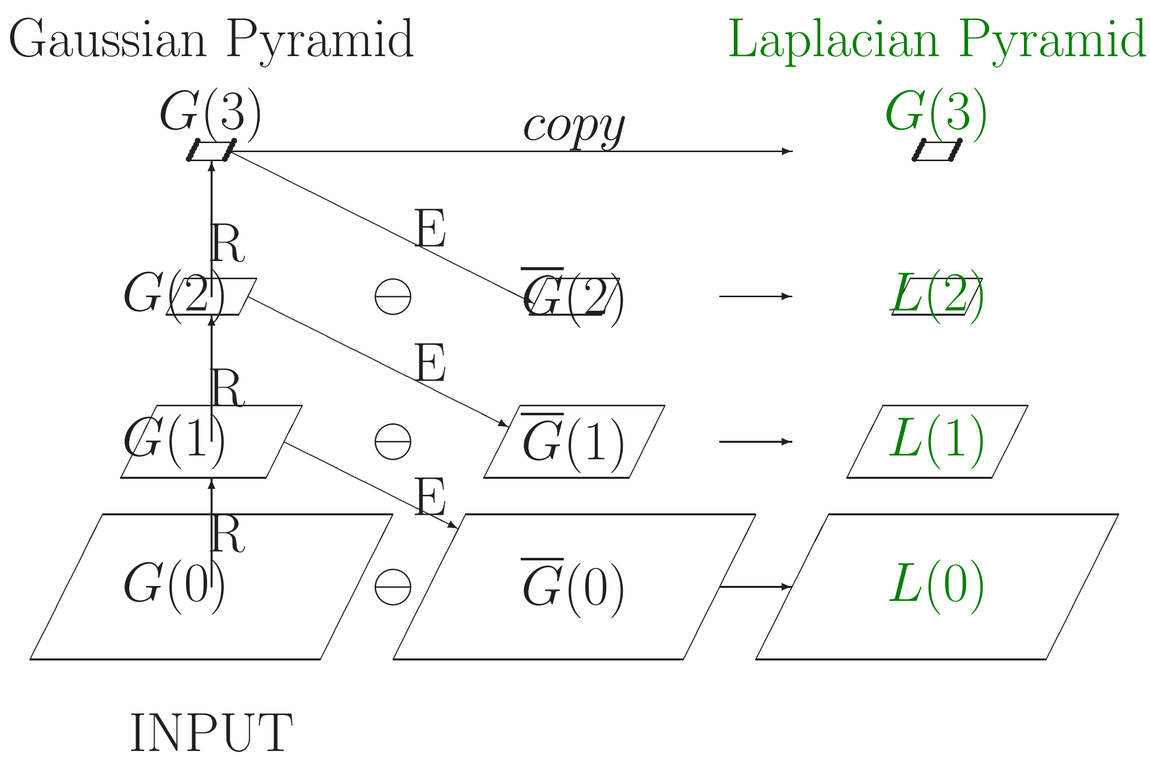
$2 \times 2/4$ Regular Image Pyramid.



a) vertical structure

b) Gaussian image pyramid

Construction of the Laplacian Pyramid

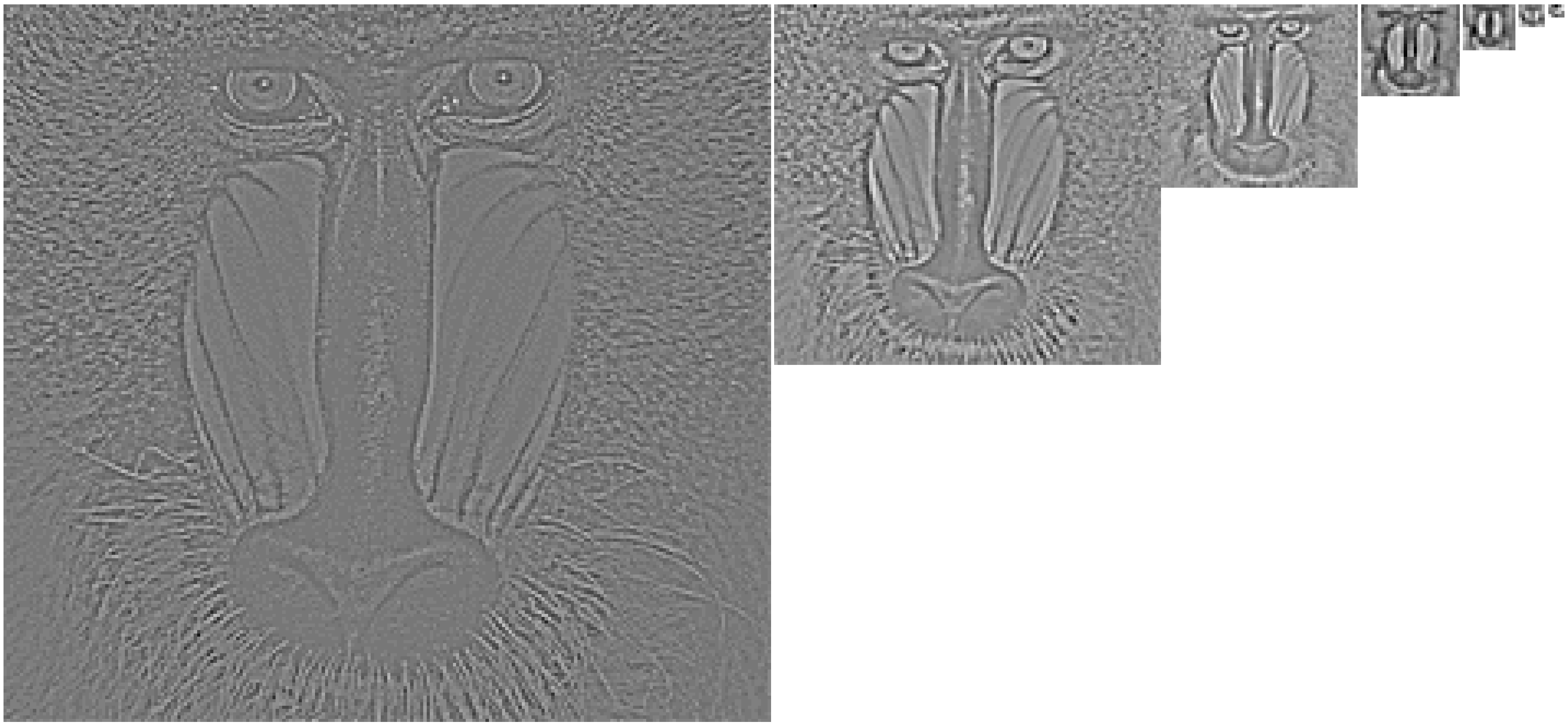


Reduction function $G(i+1) = R(G(i))$

Expansion function $\overline{G}(i-1) = E(G(i))$

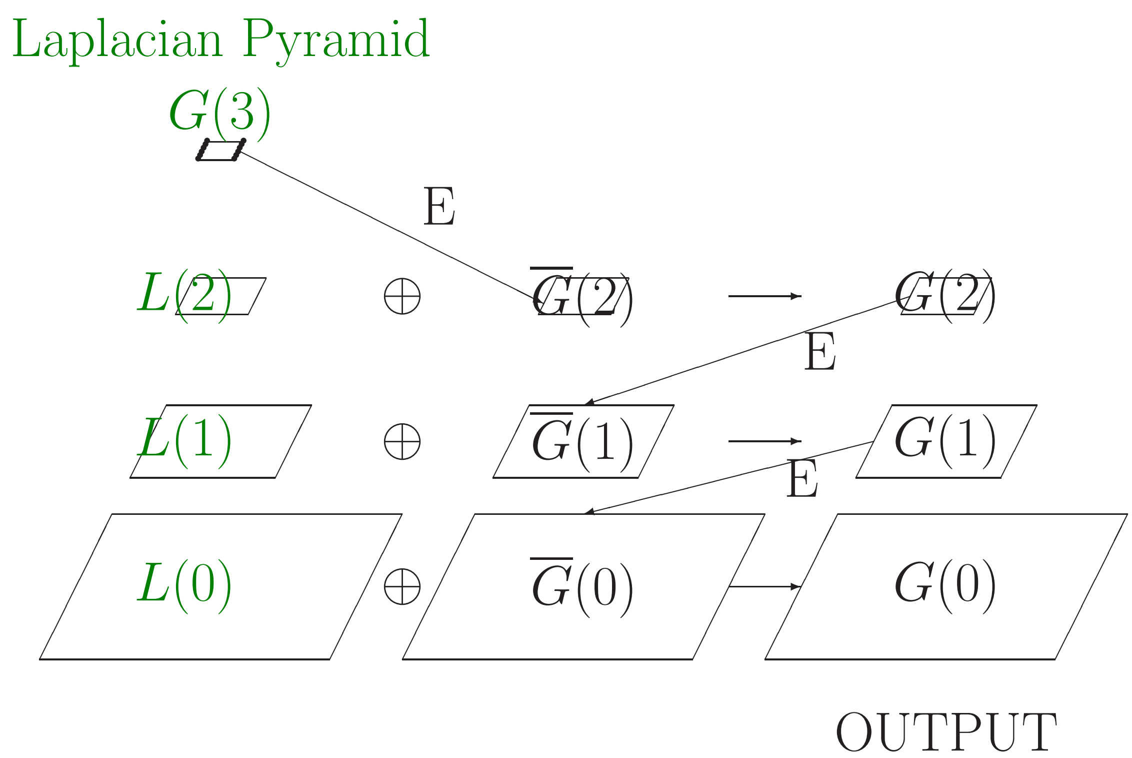
level of Laplacian $L(i) = G(i) - E(R(G(i)))$

Laplacian Pyramid



many small values (0 = grey) Quantisation Compression [BA83a]

Reconstruct Image from Laplacian Pyramid



Without Quantisation — exact Reconstruction of $G(0)$

with Quantisation — $\Longrightarrow$ 20:1 COMPRESSION

The RE-Laplacian Pyramid

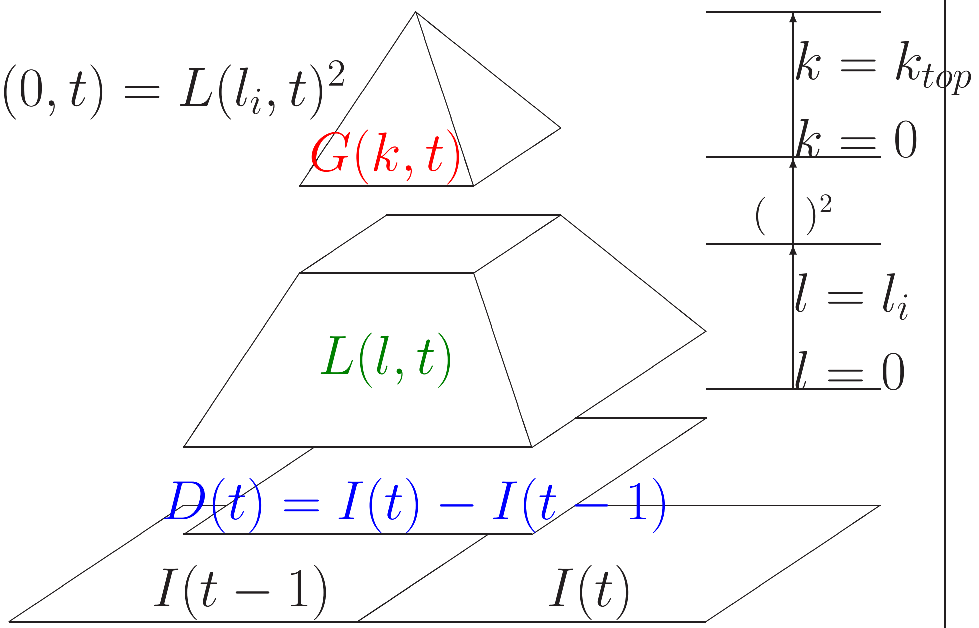
- Gaussian Pyramid $= (G_0, G_1, \ldots, G_{n-1}, G_n); G_{i+1} = R(G_i), i = 0, 1, \ldots, n-1$
- Laplacian Pyramid $= (L_0, L_1, \ldots, L_{n-1}, G_n)$ with
- EXPAND: $L_i = G_i - E(G_{i+1}) \quad i = 0, 1, \ldots, n-1$
- perfect reconstruction: $G_i = L_i + E(G_{i+1}) \quad i = n-1, n-2, \ldots, 0$

- Properties:
  - independent of reduction function R for building G
  - independent of expansion function E for building L
  - levels below the apex are invariant to brightness changes
  - brightness change of Apex $G_n \Longrightarrow$ change in the reconstructed base G_0

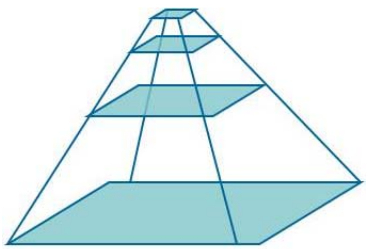
Walter G. Kropatsch TU-Wien 183/2
Pattern Recognition and Image Processing Group

1985: Change Detection and Tracking

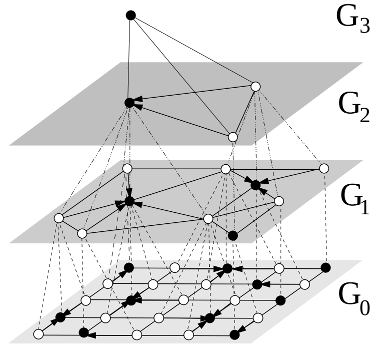
1. Frame difference $D(t) = I(t) - I(t-1)$ [ABv85]
2. $L(l,t)$ Laplace Pyramid on top of $L(0,t) = D(t)$ up to level L_i
3. Stop at level l_i = target frequency
4. $G(k,t)$ Gaussian Pyramid on top of $G(0,t) = L(l_i,t)^2$
5. Threshold $G(k_{top},t)$: Alarm



Irregular Graph PYRAMID



regular pyramid



irregular pyramid

regular pyramid		irregular pyramid
level i = image	$\longrightarrow$	graph $G_i = (V_i, E_i)$, combinatorial map, ...
implicit reduction window	$\longrightarrow$	adaptively selected contraction kernel (CK) $K_i \subset G_i$
constant reduction factor (e.g. 4)	$\longrightarrow$	$\frac{\vert V_i \vert}{\vert V_{i+1} \vert} = \frac{\vert \bigcup K_i \vert}{CC(K_i)}$
implicit level-to-level corresp.	$\longrightarrow$	contraction + simplification (dual graph contraction [KM95, Kro95])

Short recall ...

BOTTOM-UP CONSTRUCTION

Input: base level = pair of dual graphs

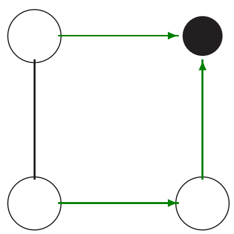
while further abstraction is possible do

1. **select contraction kernels**
2. **perform contraction**
3. **and simplification;**
4. **apply reduction functions → new reduced content**

Ouput: irregular pyramid

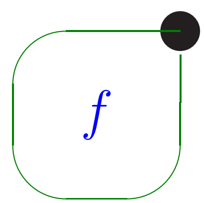
Contraction Kernel, Simplification

contract more



contraction kernel

Spanning tree



self-loop

$\deg(f) < 3$

simplified

removed

From 2D Images to 3D Dynamic Scenes

- Images are 2D samplings of 3D scenes.
- Vision sensors observe 2D projections of 3D objects in a dynamic environment.
- Volumetric 3D moving objects are covered by a **closed surface**.
- We see only (the visible part of) the surface, no interior
- Objects consist of several connected 3D parts ('constellation') and these parts can be connected in different ways:
- **rigidly, articulated, smoothly deformable.**
- In most cases objects move independently and smoothly.

Moving Structures

- Tracking a single target
- Tracking two related target points
- what is structure?
- mean shift with structure
- moving graphs/triangulations

Tracking a Single Target Point

Result = **1 Trajectory**

Tracking Two Target Points

Constellation **2 Trajectories**

Structure

- **Structure of OBJECT**

  human = body + head + 2(arm + leg)

- **Structure of CONTEXT**

  walls (restrict human's movements)

- **Structure of object IN context**

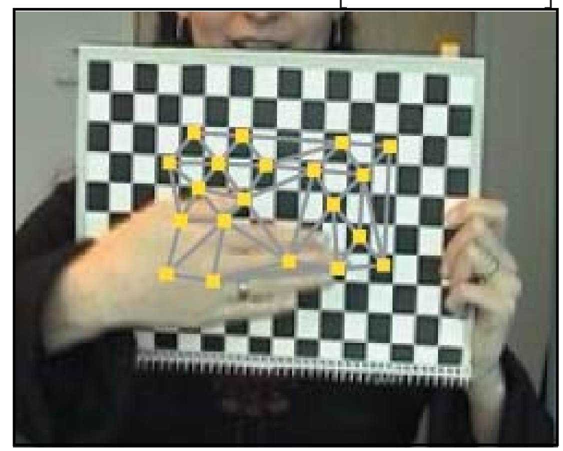
  human enters a room through a door

Mean Shift with Structure

[ALBK08]

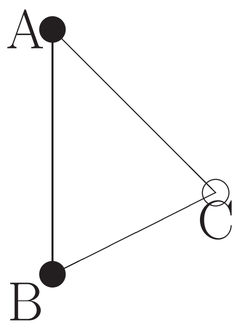
initial structure

frame 166 without



and with structure

Principle:

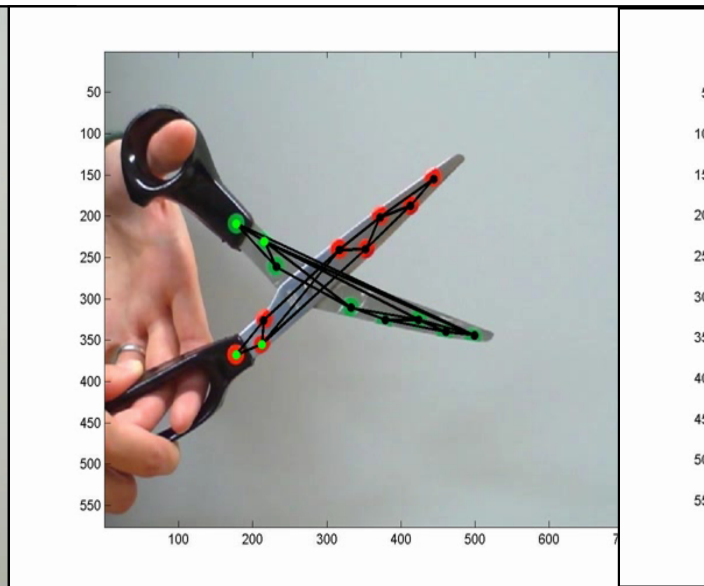


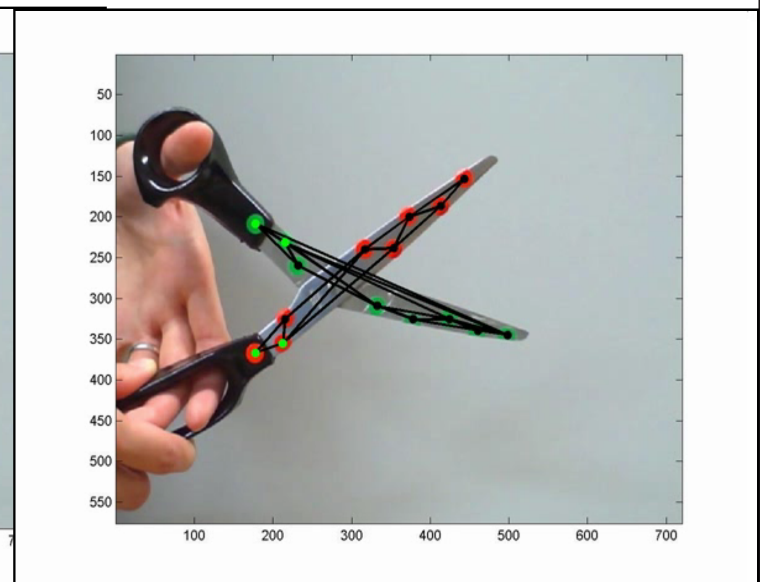
$A, B \longrightarrow C$ (by similarity of triangle)

Solution: alternate **Mean Shift** and **structure preservation**

Cutting with Scissors

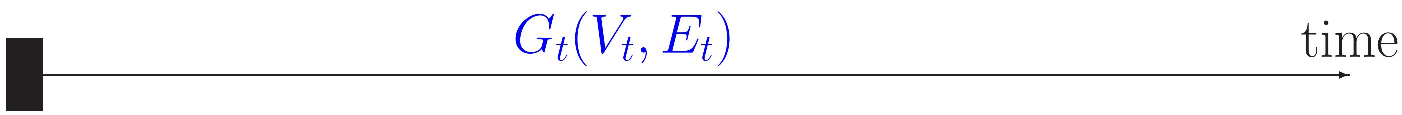
[AIK09]





initial structure tracking without and with structure

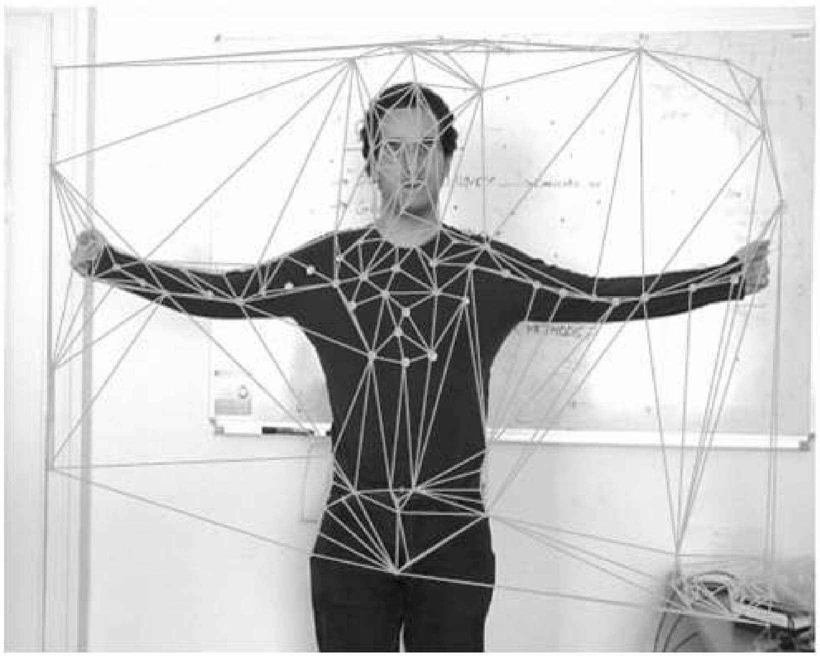
Moving Graph/Triangulation

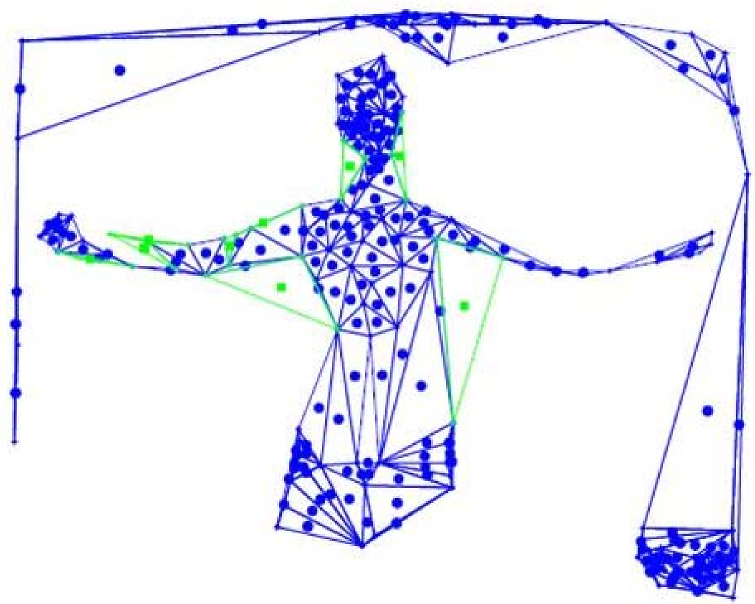
- 

- track vertices V_t of graph $G_t(V_t, E_t)$
- length of edges E_t may change or may stay the same
- triangles have 3 edges with temporal length characteristics:
  - no change: **rigid**
  - 1 edge changes: **articulation**
  - 2,3 change: **foreground/background**

Motion Segmentation

[LAI+08]



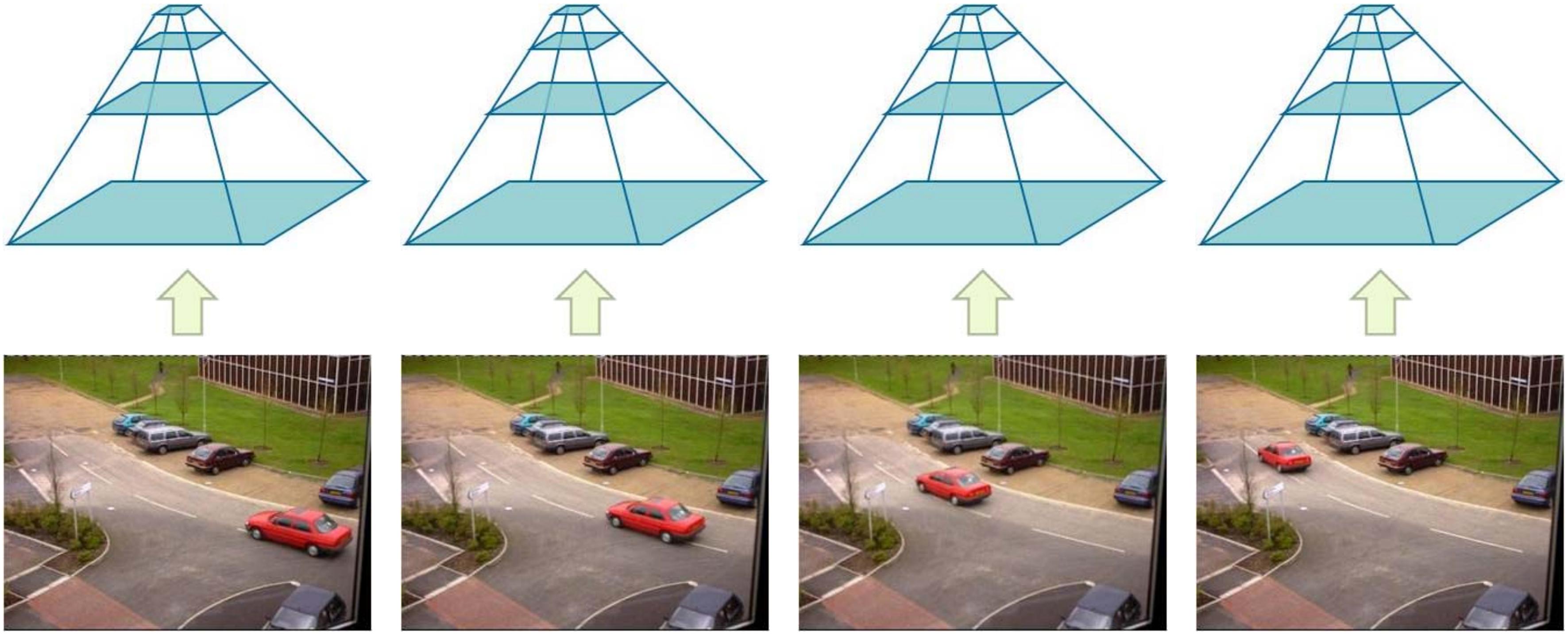


rigid articulated

TOWARDS PYRAMIDS . . .

- **one pyramid per frame**
- **one pyramid per object**
- **topological completion:**
  a pyramid on the closed object surface
- **re-introduce geometry: spatial (irregular) RE-Pyramid**

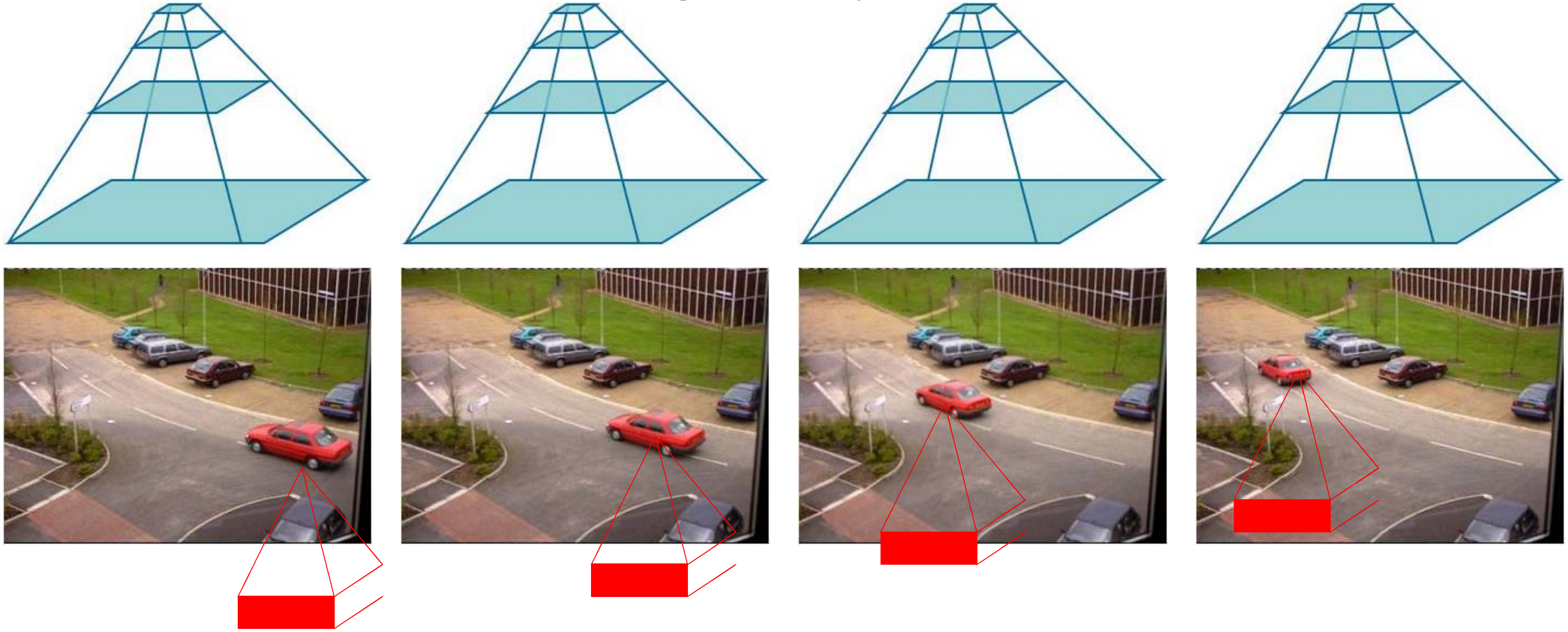
Tracking with Frame-Pyramids



Needs COMPLETE Re-Calculation: **Very time consuming**

Tracking with Moving Object-Pyramids

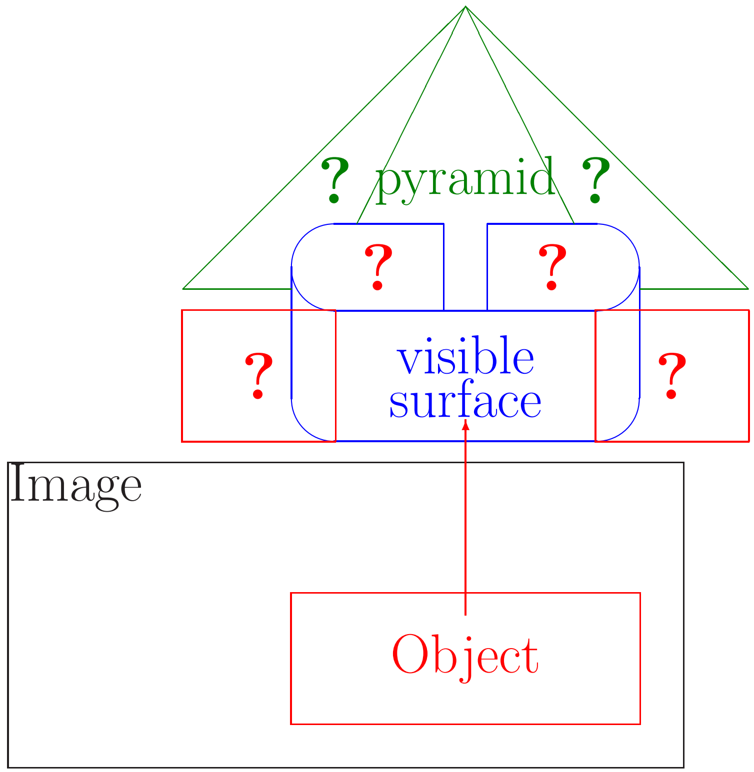
Background Pyramid



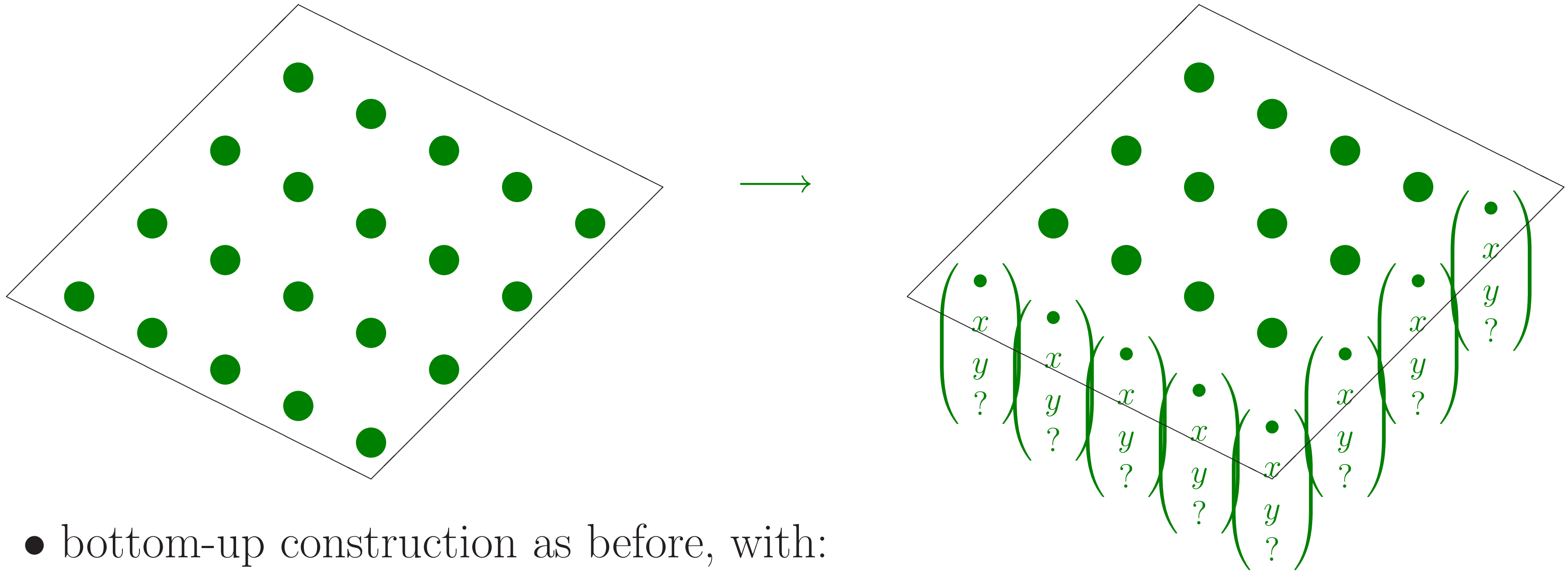
Needs Re-Calculation for object ONLY

Topological Completion

- extract **object** from image
- volumetric object is surrounded by oriented closed surface
- partly visible, partly invisible (**?**)
- build **pyramid** on
  (cyclically closed) surface
- tunnel(objects) $\longleftrightarrow$ hole(surface) $\Longrightarrow$
- plane graph $\longrightarrow$ combinatorial map

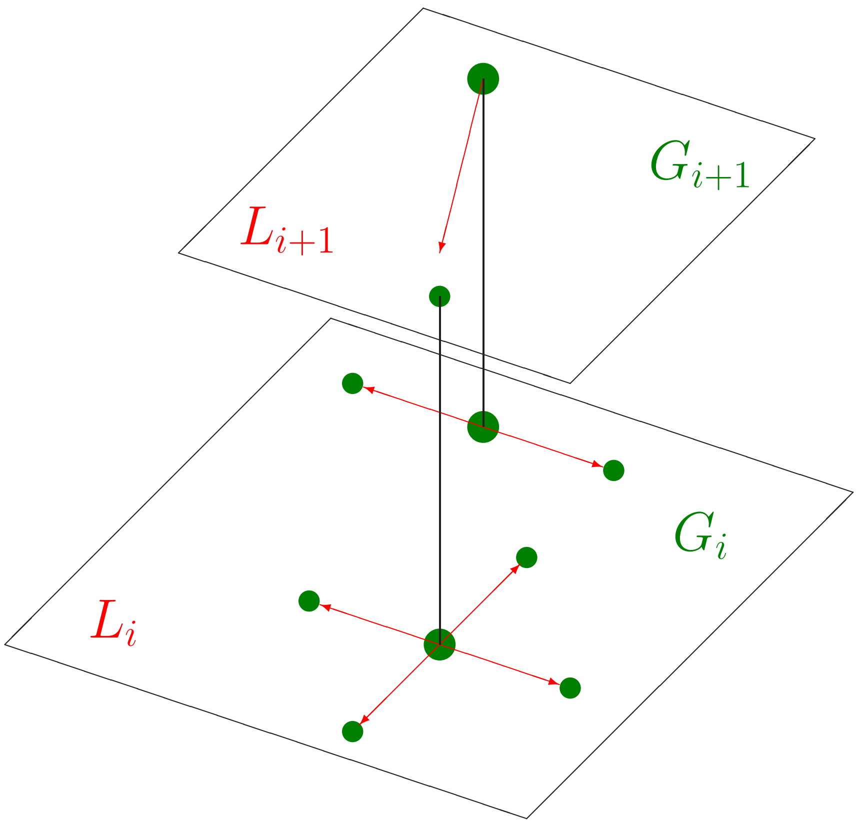


Enhance Vertices by Coordinates



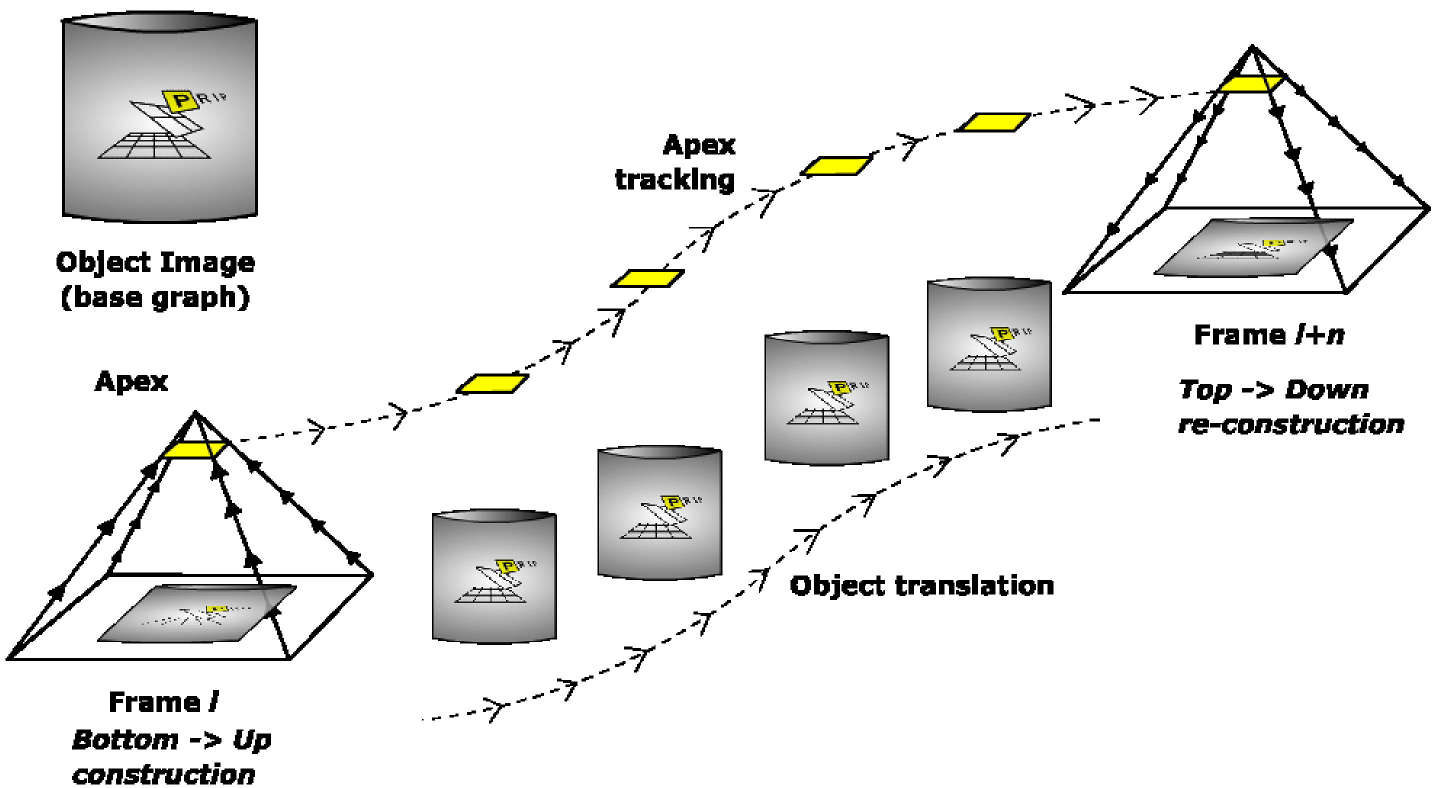
- bottom-up construction as before, with:
- content(vertex) $=$ (color, x, y, depth)$'$, position $p(c) = (x, y)'$
- simplest reduction function: inheritance, content(parent) = content(child)
- ... is as redundant as Gaussian pyramid $(G_0, G_1, \ldots, G_n)$, Laplacian?

Spatial RE-Pyramid



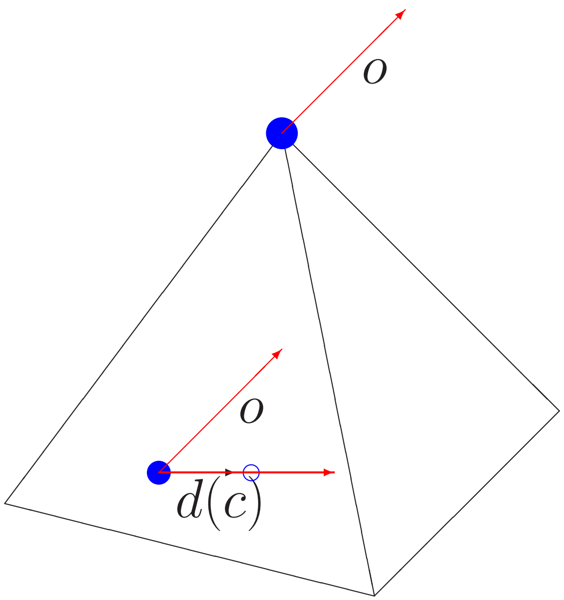
- Irregular Laplacian Pyramid:
  $L_i = G_i - E(G_{i+1})$
- correction vectors
  $d(c) = p(c) - E(p(v_p)), c \in K_i(v_p)$
- reconstruct position:
  $p(c_0) = p(c_1) + d(c_0) =$
  $= p(c_n) + \sum_{c=c_0, parent(c_0),\ldots}^{apex} d(c)$
- $L_0, L_1, \ldots L_{n-1}$ shift **invariant**

Walking Pyramid



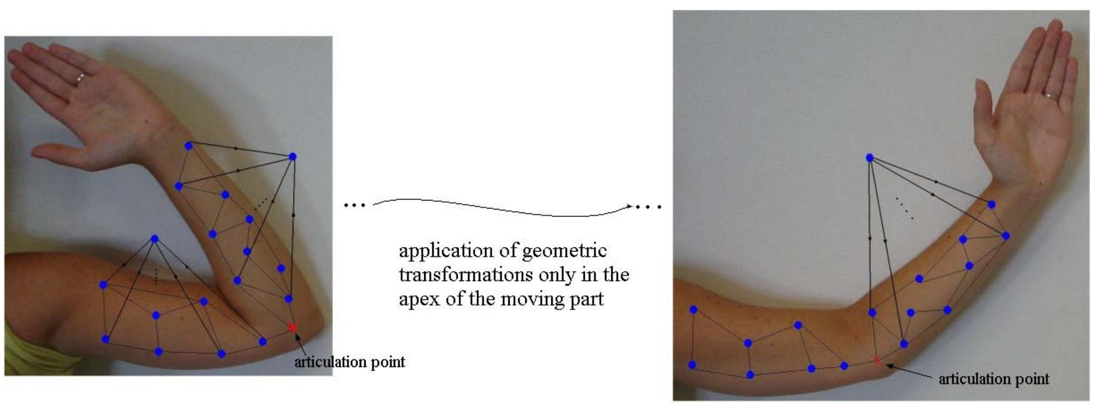
Further Invariance to Rotation and Scale

- object orientation $o \in R^3$
- $\begin{pmatrix} \bullet \\ dx \\ dy \\ dz \end{pmatrix} \longrightarrow \begin{pmatrix} \bullet \\ \lambda \\ \alpha \\ \beta \\ \gamma \end{pmatrix}$ with
- scale $\lambda = \frac{|d(c)|}{|o|}$
- and Euler angles α, β, γ to align o with $d(c)$.
- Reconstruct $d(c) = \lambda R_x(\alpha) R_y(\beta) R_z(\gamma) o$
- and position $p(c) = p(v_p) + d(c)$



Walking first Step: Articulation

Sequence with moving arm



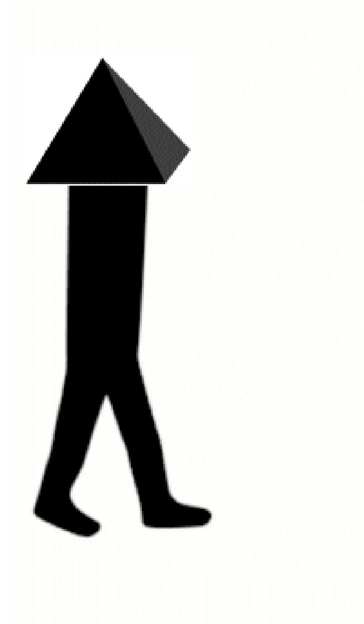
Structure initialization and Top-down reconstruction.

Conclusion

- **Moving Structure** $\longrightarrow$ **Artner CIARP09**
  - **increase robustness of tracking**
  - **largely compensate occlusions**
  - **segment F/B, rigid and articulated parts**
- **Walking Pyramids (see [Kro09])**
  - **1 pyramid per object**
  - **on surface(object) by topological completion**
  - **spatial RE-pyramid defines spatial positions**
  - **invariant to rigid movement**
  - **trajectory $\longrightarrow$ apex**

Outlook: *Running Pyramids ?*

- more sophisticated reduction functions
- multi-view integration $\longrightarrow$ **Mateos CIARP09**
- efficiency through novelty update only
- articulation
- spatio-temporal RE-pyramid ?



Thank you for your Attention

References

[AAB+84] E. H. Adelson, C. H. Anderson, J. R. Bergen, P. J. Burt, and J. M. Ogden. Pyramid methods in image processing. *RCA Engineer*, Vol. 29-6, Nov./Dec. 1984.

[ABv85] C. H. Anderson, P. J. Burt, and G. S. van der Wal. Change detection and tracking using pyramid transform techniques. *Intelligent Robots and Computer Vision*, SPIE Vol.579:pp.72–78, Sept.16-20 1985.

[AIK09] Nicole M. Artner, Adrian Ion, and Walter G. Kropatsch. Tracking Objects Beyond Rigid Motion. In Andrea Torsello, Francisco Escolano, and Mario Vento, editors, *Proceedings of the 7th IAPR-TC-15 International Workshop, GbRPR 2009, on Graph Based Representations in Pattern Recognition*, volume 5534 of *Lecture Notes in Computer Science*, pages 263–272, Berlin Heidelberg, May 2009. Springer-Verlag.

[ALBK08] Nicole M. Artner, Salvador B. López Mármol, Csaba Beleznai, and Walter G. Kropatsch. Kernel-based Tracking Using Spatial Structure. In Leila Muresan Arjan Kuipers, Bettina Heise, editor, *Challenges in the Biosciences: Image Analysis and Pattern Recognition Aspects. Proceedings of 32nd OEAGM Workshop*, pages 103 – 114, Linz, Austria, 2008. OCG-Schriftenreihe books@ocg.at, Österreichische Computer Gesellschaft. Band 232.

[BA83a] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Transactions on Communications*, Vol. COM-31(No.4):pp.532–540, April 1983.

[BA83b] P. J. Burt and E. H. Adelson. A multiresolution spline with application to image mosaics. *ACM Transactions on Graphics*, Vol. 2(No.4):pp.217–236, October 1983.

[KM95] Walter G. Kropatsch and Herwig Macho. Finding the structure of connected components using dual irregular pyramids. In *Cinquième Colloque DGCI*, pages 147–158. LLAIC1, Université d'Auvergne, ISBN 2-87663-040-0, September 1995.

[Kro95] Walter G. Kropatsch. Building Irregular Pyramids by Dual Graph Contraction. *IEE-Proc. Vision, Image and Signal Processing*, Vol. 142(No. 6):pp. 366–374, December 1995.

[Kro09] Walter G. Kropatsch. When Pyramids Learned Walking. In Jan Olof Eklundh Eduardo Bayro-Corrochano, editor, *The 14th International Congress on Pattern Recognition, CIARP 2009*, volume 5856 of *Lecture Notes in Computer Science*, pages 397–414, Berlin Heidelberg, November 2009. Springer-Verlag.

[LAI+08] Salvador B. López Mármol, Nicole Artner, Adrian Ion, Walter G. Kropatsch, and Csaba Beleznai. Video Object Segmentation Using Graphs. In José Shulcloper and Walter G. Kropatsch, editors, *Progress in Pattern Recognition, Image Analysis and Applications*, pages 733–740, Havanna, Cuba, 2008. CIARP 2008 Proceedings, Springer.