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 acknowledged!

*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:

\[
\text{HEIGHT} \leq \frac{\log(\text{image size})}{\log(\lambda)} = n.
\]

Access to $\lambda^n$ data in $n$ steps.
2 × 2/4 Regular Image Pyramid.

a) vertical structure

b) Gaussian image pyramid

Construction of the Laplacian Pyramid
Construction of the Laplacian Pyramid

Gaussian Pyramid

Laplacian Pyramid

INPUT

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

Laplacian Pyramid

\[ G(3) \]
\[ L(2) \oplus \overline{G(2)} \rightarrow G(2) \]
\[ L(1) \oplus \overline{G(1)} \rightarrow G(1) \]
\[ L(0) \oplus \overline{G(0)} \rightarrow G(0) \]

OUTPUT

Without Quantisation

exact Reconstruction of \( G(0) \) \( \Rightarrow 20:1 \) COMPRESSION

with Quantisation
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: \(L_i = G_i - E(G_{i+1})\)  \(i = 0, 1, \ldots, n-1\)
• perfect reconstruction: \(G_i = L_i + E(G_{i+1})\)  \(i = n-1, n-2, \ldots, 0\)

• 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 \(G_n\) \(\Rightarrow\) change in the reconstructed base \(G_0\)

1985: Change Detection and Tracking
1985: Change Detection and Tracking

1. Frame difference \( D(t) = I(t) - I(t - 1) \) \[\text{[ABv85]}\]

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

level $i = \text{image}$ $\rightarrow$ graph $G_i = (V_i, E_i)$, combinatorial map, ...
implicit reduction window $\rightarrow$ adaptively selected
constant reduction factor (e.g. 4) $\rightarrow \frac{|V_i|}{|V_{i+1}|} = \frac{|\bigcup K_i|}{CC(K_i)}$
implicit level-to-level corresp. $\rightarrow$ 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

Output: irregular pyramid
Contraction Kernel, Simplification

contract more

contraction kernel

self-loop

Spanning tree

deg(f) < 3

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

2 Trajectories
Structure

• Structure of OBJECT
  \[ \text{human} = \text{body} + \text{head} + 2(\text{arm} + \text{leg}) \]

• Structure of CONTEXT
  \[ \text{walls} \text{ (restrict human’s movements)} \]

• Structure of object IN context
  human enters a room through a door
Mean Shift with Structure

initial structure

frame 166 without

and with structure

Principle: \( A, B \rightarrow C \) (by similarity of triangle)

Solution: alternate Mean Shift and structure preservation
Cutting with Scissors

[AIK09]
Moving Graph/Triangulation

\[ G_t(V_t, E_t) \]

- 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**
Walking Pyramids

Motion Segmentation

rigid
articulated

[LAI+08]
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) $\leftrightarrow$ hole(surface) $\Rightarrow$
- plane graph $\longrightarrow$ combinatorial map

Image

Object

Enhance Vertices by Coordinates
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 (G₀, G₁, ..., Gₙ), 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:
  \[ p(c_0) = p(c_1) + d(c_0) = p(c_n) + \sum_{c=\text{c}_0, \text{parent}(c_0),...} d(c) \]

- \( L_0, L_1, \ldots L_{n-1} \) shift invariant
Walking Pyramids

Walking Pyramid

Object Image (base graph)

Apex

Frame i
Bottom -> Up construction

Frame i+n
Top -> Down re-construction

Apex tracking

Object translation

Further Invariance to Rotation and Scale
Further Invariance to Rotation and Scale

- object orientation $o \in \mathbb{R}^3$

$$
\begin{pmatrix}
\cdot \\
\cdot \\
\cdot \\
dx \\
dy \\
dz
\end{pmatrix}
\rightarrow
\begin{pmatrix}
\cdot \\
\cdot \\
\cdot \\
\lambda \\
\alpha \\
\beta \\
\gamma
\end{pmatrix}
$$

with

- scale $\lambda = \frac{|d(c)|}{|o|}$
- and Euler angles $\alpha, \beta, \gamma$ 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
  - 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 → apex

Artner CIARP09

Outlook: Running Pyramids ?
Outlook: *Running Pyramids*?

- more sophisticated reduction functions
- multi-view integration
- efficiency through novelty update only
- articulation
- spatio-temporal RE-pyramid?

Thank you for your Attention
References


Walking Pyramids