

Pyramid (image processing)

Pyramid, or **pyramid representation**, is a type of multi-scale signal representation developed by the computer vision, image processing and signal processing communities, in which a signal or an image is subject to repeated smoothing and subsampling. Pyramid representation is a predecessor to scale-space representation and multiresolution analysis.

Contents

Pyramid generation

Pyramid generation kernels

Gaussian pyramid

Laplacian pyramid

Steerable pyramid

Applications of pyramids

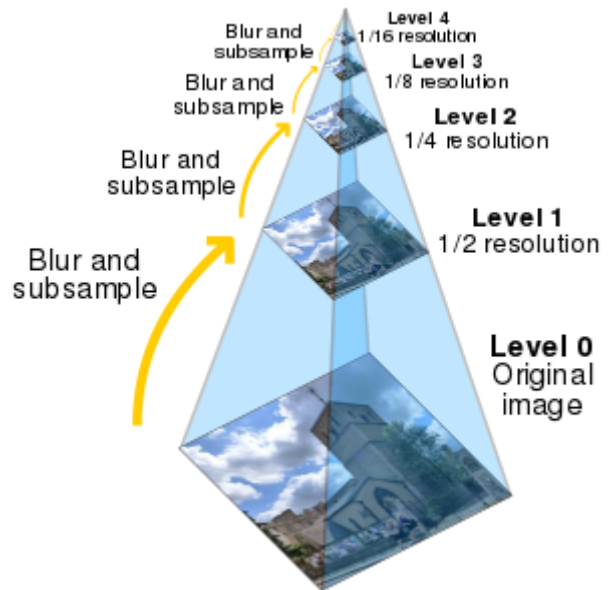
Alternative representation

Detail manipulation

See also

References

External links



Visual representation of an image pyramid with 5 levels

Pyramid generation

There are two main types of pyramids: lowpass and bandpass.

A lowpass pyramid is made by smoothing the image with an appropriate smoothing filter and then subsampling the smoothed image, usually by a factor of 2 along each coordinate direction. The resulting image is then subjected to the same procedure, and the cycle is repeated multiple times. Each cycle of this process results in a smaller image with increased smoothing, but with decreased spatial sampling density (that is, decreased image resolution). If illustrated graphically, the entire multi-scale representation will look like a pyramid, with the original image on the bottom and each cycle's resulting smaller image stacked one atop the other.

A bandpass pyramid is made by forming the difference between images at adjacent levels in the pyramid and performing image interpolation between adjacent levels of resolution, to enable computation of pixelwise differences.^[1]

Pyramid generation kernels

A variety of different smoothing kernels have been proposed for generating pyramids.^{[2][3][4][5][6][7]} Among the suggestions that have been given, the *binomial kernels* arising from the binomial coefficients stand out as a particularly useful and theoretically well-founded class.^{[3][8][9][10]} Thus, given a two-dimensional image, we may apply the (normalized) binomial filter (1/4, 1/2, 1/4) typically twice or more along each spatial dimension and then subsample the image by a factor of two. This operation may then proceed as many times as desired, leading to a compact and efficient multi-scale representation. If motivated by specific requirements, intermediate scale levels may also be generated where the subsampling stage is sometimes left out, leading to an *oversampled* or *hybrid pyramid*.^[11] With the increasing computational efficiency of CPUs available today, it is in some situations also feasible to use wider support Gaussian filters as smoothing kernels in the pyramid generation steps.

Gaussian pyramid

In a Gaussian pyramid, subsequent images are weighted down using a Gaussian average (Gaussian blur) and scaled down. Each pixel containing a local average corresponds to a neighborhood pixel on a lower level of the pyramid. This technique is used especially in texture synthesis.

Laplacian pyramid

A Laplacian pyramid is very similar to a Gaussian pyramid but saves the difference image of the blurred versions between each levels. Only the smallest level is not a difference image to enable reconstruction of the high resolution image using the difference images on higher levels. This technique can be used in image compression.^[12]

Steerable pyramid

A steerable pyramid, developed by Simoncelli and others, is an implementation of a multi-scale, multi-orientation band-pass filter bank used for applications including image compression, texture synthesis, and object recognition. It can be thought of as an orientation selective version of a Laplacian pyramid, in which a bank of steerable filters are used at each level of the pyramid instead of a single Laplacian or Gaussian filter.^{[13][14][15]}

Applications of pyramids

Alternative representation

In the early days of computer vision, pyramids were used as the main type of multi-scale representation for computing multi-scale image features from real-world image data. More recent techniques include scale-space representation, which has been popular among some researchers due to its theoretical foundation, the ability to decouple the subsampling stage from the multi-scale representation, the more powerful tools for theoretical analysis as well as the ability to compute a representation at *any* desired scale, thus avoiding the algorithmic problems of relating image representations at different resolution. Nevertheless, pyramids are still frequently used for expressing computationally efficient approximations to scale-space representation.^{[11][16][17]}

Detail manipulation

Laplacian image pyramids based on the bilateral filter provide a good framework for image detail enhancement and manipulation. The difference images between each layer are modified to exaggerate or reduce details at different scales in an image.

Some image compression file formats use the Adam7 algorithm or some other interlacing technique. These can be seen as a kind of image pyramid. Because those file format store the "large-scale" features first, and fine-grain details later in the file, a particular viewer displaying a small "thumbnail" or on a small screen can quickly download just enough of the image to display it in the available pixels—so one file can support many viewer resolutions, rather than having to store or generate a different file for each resolution.

See also

- Mipmap
- Scale space implementation
- Level of detail
- JPEG 2000#Multiple resolution representation

References

1. E.H. Adelson and C.H. Anderson and J.R. Bergen and P.J. Burt and J.M. Ogden. "Pyramid methods in image processing" (http://persci.mit.edu/pub_pdfs/RCA84.pdf). 1984.
2. Burt, P. J. (May 1981). "Fast filter transform for image processing". *Computer Graphics and Image Processing*. **16**: 20–51. doi:10.1016/0146-664X(81)90092-7 (<https://doi.org/10.1016%2F0146-664X%2881%2990092-7>).
3. Crowley, James L. (November 1981). "A representation for visual information" (http://www.ri.cmu.edu/publication_view.html?pub_id=37). Carnegie-Mellon University, Robotics Institute. tech. report CMU-RI-TR-82-07.
4. Burt, Peter; Adelson, Ted (1983). "The Laplacian Pyramid as a Compact Image Code" (http://persci.mit.edu/pub_pdfs/pyramid83.pdf) (PDF). *IEEE Trans. Commun.* **9** (4): 532–540. CiteSeerX 10.1.1.54.299 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.299>). doi:10.1109/TCOM.1983.1095851 (<https://doi.org/10.1109%2FTCOM.1983.1095851>).
5. Crowley, J. L.; Parker, A. C. (March 1984). "A representation for shape based on peaks and ridges in the difference of low-pass transform". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **6** (2): 156–170. CiteSeerX 10.1.1.161.3102 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.161.3102>). doi:10.1109/TPAMI.1984.4767500 (<https://doi.org/10.1109%2FTPAMI.1984.4767500>). PMID 21869180 (<https://pubmed.ncbi.nlm.nih.gov/21869180>).
6. Crowley, J. L.; Sanderson, A. C. (1987). "Multiple resolution representation and probabilistic matching of 2-D gray-scale shape" (<http://www-prima.inrialpes.fr/Prima/Homepages/jlc/papers/Crowley-Sanderson-PAMI87.pdf>) (PDF). *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **9** (1): 113–121. CiteSeerX 10.1.1.1015.9294 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.1015.9294>). doi:10.1109/tpami.1987.4767876 (<https://doi.org/10.1109%2FTpami.1987.4767876>). PMID 21869381 (<https://pubmed.ncbi.nlm.nih.gov/21869381>).
7. Meer, P.; Baugher, E. S.; Rosenfeld, A. (1987). "Frequency domain analysis and synthesis of image generating kernels". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. **9** (4): 512–522. doi:10.1109/tpami.1987.4767939 (<https://doi.org/10.1109%2FTpami.1987.4767939>). PMID 21869409 (<https://pubmed.ncbi.nlm.nih.gov/21869409>).
8. Lindeberg, Tony, "Scale-space for discrete signals" (<http://www.nada.kth.se/~tony/abstracts/Lin90-PAMI.html>), "PAMI(12), No. 3, March 1990, pp. 234-254.

9. Lindeberg, Tony. Scale-Space Theory in Computer Vision (<http://www.nada.kth.se/~tony/book.html>), Kluwer Academic Publishers, 1994, ISBN 0-7923-9418-6 (see specifically Chapter 2 for an overview of Gaussian and Laplacian image pyramids and Chapter 3 for theory about generalized binomial kernels and discrete Gaussian kernels)
10. See the article on multi-scale approaches for a very brief theoretical statement
11. Lindeberg, T. and Bretzner, L. Real-time scale selection in hybrid multi-scale representations (<http://www.nada.kth.se/cvap/abstracts/cvap279.html>), Proc. Scale-Space'03, Isle of Skye, Scotland, Springer Lecture Notes in Computer Science, volume 2695, pages 148-163, 2003.
12. Burt, Peter J.; Adelson, Edward H. (1983). "The Laplacian Pyramid as a Compact Image Code" (http://persci.mit.edu/pub_pdfs/pyramid83.pdf) (PDF). *IEEE Transactions on Communications*. **31** (4): 532–540. CiteSeerX 10.1.1.54.299 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.54.299>). doi:10.1109/TCOM.1983.1095851 (<https://doi.org/10.1109%2FTCOM.1983.1095851>).
13. Simoncelli, Eero. "The Steerable Pyramid" (<http://www.cns.nyu.edu/~eero/STEERPYPYR/>). cns.nyu.edu.
14. Manduchi, Roberto; Perona, Pietro; Shy, Doug (1997). "Efficient Deformable Filter Banks" (http://www.vision.caltech.edu/publications/ManduchiPeronaShy_efficient_deformable.pdf) (PDF). California Institute of Technology/University of Padua.
Also in Manduchi, R.; Perona, P.; Shy, D. (1998). "Efficient Deformable Filter Banks". *Transactions on Signal Processing*. **46** (4): 1168–1173. Bibcode:1998ITSP...46.1168M (<https://ui.adsabs.harvard.edu/abs/1998ITSP...46.1168M>). CiteSeerX 10.1.1.5.3102 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.5.3102>). doi:10.1109/78.668570 (<https://doi.org/10.1109%2F78.668570>).
15. Stanley A. Klein ; Thom Carney ; Lauren Barghout-Stein and Christopher W. Tyler "Seven models of masking", Proc. SPIE 3016, Human Vision and Electronic Imaging II, 13 (June 3, 1997); doi:10.1117/12.274510 (<https://doi.org/10.1117%2F12.274510>)
16. Crowley, J, Riff O. Fast computation of scale normalised Gaussian receptive fields (<http://www-prima.inrialpes.fr/Prima/Homepages/jlc/papers/Crowley-ScaleSpace03.pdf>), Proc. Scale-Space'03, Isle of Skye, Scotland, Springer Lecture Notes in Computer Science, volume 2695, 2003.
17. Lowe, D. G. (2004). "Distinctive image features from scale-invariant keypoints" (<http://citeseer.ist.psu.edu/lowe04distinctive.html>). *International Journal of Computer Vision*. **60** (2): 91–110. CiteSeerX 10.1.1.73.2924 (<https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.73.2924>). doi:10.1023/B:VISI.0000029664.99615.94 (<https://doi.org/10.1023%2FB%3AVISI.0000029664.99615.94>).

External links

- Gaussian-Laplacian Pyramid Image Coding (<http://fourier.eng.hmc.edu/e161/lectures/canny/node3.html>) - illustrates methods of Downsampling, Upsampling, and Gaussian convolution
- The Gaussian Pyramid (http://www.cse.yorku.ca/~kosta/CompVis_Notes/gaussian_pyramid.pdf) - provides a brief introduction for the procedure and cites several sources
- Laplacian Irregular Graph Pyramid (<http://www.prip.tuwien.ac.at/twist/irregulargraphpyramid.php>) - Figure 1 on this page illustrates an example of the Gaussian Pyramid
- The Laplacian Pyramid as a Compact Image Code (<https://archive.is/20130117073108/http://aspdf.com/ebook/the-laplacian-pyramid-as-a-compact-image-code-pdf.html>) on eBook Submission (<http://aspdf.com/>)

Retrieved from "[https://en.wikipedia.org/w/index.php?title=Pyramid_\(image_processing\)&oldid=993850807](https://en.wikipedia.org/w/index.php?title=Pyramid_(image_processing)&oldid=993850807)"

Text is available under the Creative Commons Attribution-ShareAlike License; additional terms may apply. By using this site, you agree to the Terms of Use and Privacy Policy. Wikipedia® is a registered trademark of the Wikimedia Foundation, Inc., a non-profit organization.