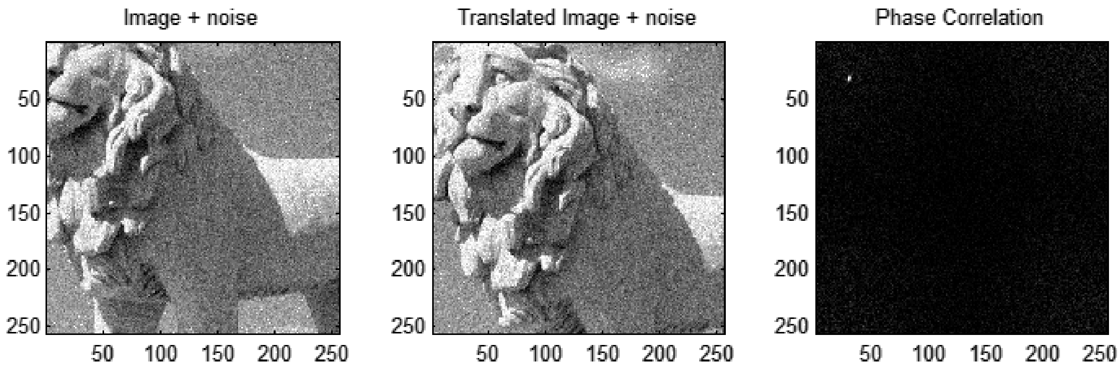
Phase correlation

In image processing, **phase correlation** is a method of image registration, and uses a fast frequency-domain approach to estimate the relative translative offset between two similar images.

Example

The following image demonstrates the usage of phase correlation to determine relative translative movement between two images corrupted by independent Gaussian noise. The image was translated by (30,33) pixels. Accordingly, one can clearly see a peak in the phase-correlation representation at approximately (30,33).



Method

Given two input images g_a and g_b:

Apply a window function (e.g., a Hamming window) on both images to reduce edge effects. Then, calculate the discrete 2D Fourier transform of both images.

$$\mathbf{G}_a = \mathcal{F}\{g_a\}, \ \mathbf{G}_b = \mathcal{F}\{g_b\}$$

Calculate the cross-power spectrum by taking the complex conjugate of the second result, multiplying the Fourier transforms together elementwise, and normalizing this product elementwise.

$$R = \frac{\mathbf{G}_a \mathbf{G}_b^*}{|\mathbf{G}_a \mathbf{G}_b^*|}$$

Obtain the normalized cross-correlation by applying the inverse Fourier transform.

$$r = \mathcal{F}^{-1}\{R\}$$

Determine the location of the peak in r (possibly using sub-pixel edge detection).

$$(\Delta x, \Delta y) = \arg\max_{(x,y)}\{r\}$$

Rationale

The method is based on the Fourier shift theorem. Let the two images g_a and g_b be circularly-shifted versions of each other:

$$g_b(x,y) \stackrel{\text{def}}{=} g_a((x-\Delta x) \bmod M, (y-\Delta y) \bmod N)$$

(where the images are $M \times N$ in size).

Then, the discrete Fourier transforms of the images will be shifted relatively in phase:

$$\mathbf{G}_b(u,v) = \mathbf{G}_a(u,v)e^{-2\pi i(\frac{u\Delta x}{M}+\frac{v\Delta y}{N})}$$

One can then calculate the normalized cross-power spectrum to factor out the phase difference:

$$\begin{aligned} R(u,v) &= \frac{\mathbf{G}_a\mathbf{G}_b^*}{|\mathbf{G}_a\mathbf{G}_b^*|} \\ &= \frac{\mathbf{G}_a\mathbf{G}_a^* e^{2\pi i(\frac{u\Delta x}{M}+\frac{v\Delta y}{N})}}{|\mathbf{G}_a\mathbf{G}_a^* e^{2\pi i(\frac{u\Delta x}{M}+\frac{v\Delta y}{N})}|} \\ &= \frac{\mathbf{G}_a\mathbf{G}_a^* e^{2\pi i(\frac{u\Delta x}{M}+\frac{v\Delta y}{N})}}{|\mathbf{G}_a\mathbf{G}_a^*|} \\ &= e^{2\pi i(\frac{u\Delta x}{M}+\frac{v\Delta y}{N})} \end{aligned}$$

since the magnitude of a complex exponential always is one, and the phase of $\mathbf{G}_a\mathbf{G}_a^*$ always is zero.

The inverse Fourier transform of a complex exponential is a Kronecker delta, i.e. a single peak:

$$r(x,y) = \delta(x+\Delta x, y+\Delta y)$$

This result could have been obtained by calculating the cross correlation directly. The advantage of this method is that the discrete Fourier transform and its inverse can be performed using the fast Fourier transform, which is much faster than correlation for large images.

Benefits

Unlike many spatial-domain algorithms, the phase correlation method is resilient to noise, occlusions, and other defects typical of medical or satellite images.

The method can be extended to determine rotation and scaling differences between two images by first converting the images to log-polar coordinates. Due to properties of the Fourier transform, the rotation and scaling parameters can be determined in a manner invariant to translation.[1] [2]

Limitations

In practice, it is more likely that g_b will be a simple linear shift of g_a, rather than a circular shift as required by the explanation above. In such cases, r will not be a simple delta function, which will reduce the performance of the method. In such cases, a window function should be employed during the Fourier transform to reduce edge effects, or the images should be zero padded so that the edge effects can be ignored. If the images consist of a flat background, with all detail situated away from the edges, then a linear shift will be equivalent to a circular shift, and the above derivation will hold exactly.

For periodic images (such as a chessboard), phase correlation may yield ambiguous results with several peaks in the resulting output.

Applications

Phase correlation is the most preferred method for television systems conversion, as it leaves the fewest artefacts.

TV Systems Conversion

Phase Correlation is perhaps the most computationally complex of the general algorithms used in television systems conversion.

Phase Correlation's success lies in the fact that it is effective with coping with rapid motion and random motion. Phase Correlation doesn't easily get confused by rotating or twirling objects that confuse most other kinds of systems converters.

Phase Correlation is elegant as well as technically and conceptually complex. Its successful operation is derived by performing a Fourier Transform to each field of video.

A Fast Fourier Transform (FFT) is an algorithm which deals with the transformation of discrete values (in this case image pixels).

When applied to a sample of finite values, a Fast Fourier Transform expresses any changes (motion) in terms of frequency components.

What is the advantage of using FFTs over simply trying to predict the motion vector on a pixel by pixel basis?

- Mathematically, it's far easier and faster to recognize and process frequency signatures from which very accurate motion vectors can then be calculated.
- *Rather than having to measure where every pixel goes from frame to frame the FFT rather results in representing just the changes from one frame to the next.*

Since the result of the FFT represents only the inter-frame changes in terms of frequency distribution, there's far less data that has to be processed in order to calculate the motion vectors.

- Unlike other motion vector calculating methods, the FFT technique is not easily fooled by objects that have rotational or spiraling motions.
- What results from the FFT is a three dimensional frequency distribution represented mathematically by peaks in a three dimensional wave pattern.
- The 3rd dimension in this coordinate system represents subsequent fields of video.

In summation: Objects in motion can be mathematically correlated to their peaks in the frequency distribution. Once the FFT is performed it becomes a computationally simple matter for the computer to track just the peaks and assign them the appropriate motion vectors. This conversion technique is both elegant and computationally involved. Sophisticated software and large amounts of processor "horsepower" are required for these complex computations.

See also

General

- Cross correlation

Television

- Television standards conversion
- Reverse Standards Conversion

External links

- Using Matlab to perform phase correlation on images [3]

References

[1] E. De Castro and C. Morandi " Registration of Translated and Rotated Images Using Finite Fourier Transforms (http://www.idi.ntnu.no/~fredrior/files/Registration of Translated and Rotated Images Using FT.pdf)", IEEE Transactions on pattern analysis and machine intelligence, Sept. 1987

[2] B. S Reddy and B. N. Chatterji, " An FFT-based technique for translation, rotation, and scale-invariant image registration (http://calhau.dca.fee.unicamp.br/wiki/images/3/3d/Fft_correlation.pdf)", IEEE Transactions on Image Processing 5, no. 8 (1996): 1266–1271.

[3] http://www.mathworks.com/products/demos/image/cross_correlation/imreg.html

Article Sources and Contributors

Phase correlation *Source*: http://en.wikipedia.org/w/index.php?oldid=382628704 *Contributors*: BenFrantzDale, Billlion, Chadernook, Charles Matthews, David Wahler, ENeville, Eyreland, Oli Filth, Orderud, Parapunter, SebastianHelm, Vossman, Xezbeth, 24 anonymous edits

Image Sources, Licenses and Contributors

Image:Phase correlation.png *Source*: http://en.wikipedia.org/w/index.php?title=File:Phase_correlation.png *License*: Public Domain *Contributors*: User:Orderud

License