Imperial College London

ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION

Georgios Tzimiropoulos and Tania Stathaki

"An Adaptive Intelligent Platform for Object Recognition", sponsored by SELEX

July 2009

FFT-BASED IMAGE REGISTRATION

OUR APPROACH TO DTC DEMANDS AND NEEDS

A framework based on frequency domain correlation schemes

APPLICATIONS

- Global scene representation (e.g. video frame registration)
- Image super-resolution
- Video coding
- Fast dense motion estimation (i.e. an alternative to optical flow algorithms)
- Motion-based segmentation
- Structure from motion
- Tracking
- Scale-invariant image registration (extension to affine still open question)
- Scale-invariant object detection

FFT-BASED IMAGE REGISTRATION

ADVANTAGES

- Spans a wide range of applications
- Computational efficiency
  1. Large amount of work available for FFT-based processing of data
  2. Recently proposed machine-specific optimized FFT implementations require only a few ms for a 512x512 FFT → real time registration
  3. Parallel processing grows rapidly
- Constant time of processing
- Good accuracy
- Requires the fine tuning of very few parameters
- Framework for 2D/3D/ND registration

FFT-BASED IMAGE REGISTRATION

STATE-OF-THE-ART   PHASE CORRELATION

1. Tailors correlation to the nature of images
2. Excellent peak localization accuracy
3. Immunity to illumination changes, insensitivity to changes in spectral energy

PROBLEMS WITH PHASE CORRELATION

1. Performance deteriorates in the presence of non-overlapping regions
2. Performance deteriorates in the presence of AWGN
3. Problems with border effects and aliasing

OUR APPROACH   GRADIENT CORRELATION SCHEMES

1. Extract image gradients
2. Combine them in a complex gray level edge map which retains both magnitude and orientation information
3. Correlate

ROBUST FFT-BASED IMAGE REGISTRATION

NORMALIZED GRADIENT CORRELATION

$$I_1(\mathbf{x}+\mathbf{t}) = I_2(\mathbf{x})$$

We introduce

$$\mathrm{NGC}(\mathbf{u}) \triangleq \frac{G_1(\mathbf{u}) \star G_2^*(-\mathbf{u})}{|G_1(\mathbf{u})| \star |G_2(-\mathbf{u})|} = \frac{\int_{\mathcal{R}^2} G_1(\mathbf{x}) G_2^*(\mathbf{x}+\mathbf{u}) d\mathbf{x}}{\int_{\mathcal{R}^2} |G_1(\mathbf{x})||G_2(\mathbf{x}+\mathbf{u})| d\mathbf{x}}$$

where

$$G_i(\mathbf{x}) = G_{i,x}(\mathbf{x}) + jG_{i,y}(\mathbf{x})$$

Finally

$$\widehat{\mathbf{t}} = \arg_{\mathbf{u}} \max\{\mathrm{NGC}(\mathbf{u})\}$$

ROBUST FFT-BASED IMAGE REGISTRATION

NORMALIZED GRADIENT CORRELATION   A CLOSER LOOK (1)

$$\mathrm{NGC}(\mathbf{u}) \triangleq \frac{\int_{\mathcal{R}^2} R_1(\mathbf{x}) R_2(\mathbf{x}+\mathbf{u}) \cos[\Phi_1(\mathbf{x}) - \Phi_2(\mathbf{x}+\mathbf{u})] d\mathbf{x}}{\int_{\mathcal{R}^2} R_1(\mathbf{x}) R_2(\mathbf{x}+\mathbf{u}) d\mathbf{x}}$$

where

$$R_i = \sqrt{G_{i,x}^2 + G_{i,y}^2} \qquad \Phi_i = \arctan G_{i,y}/G_{i,x}$$

- Denominator

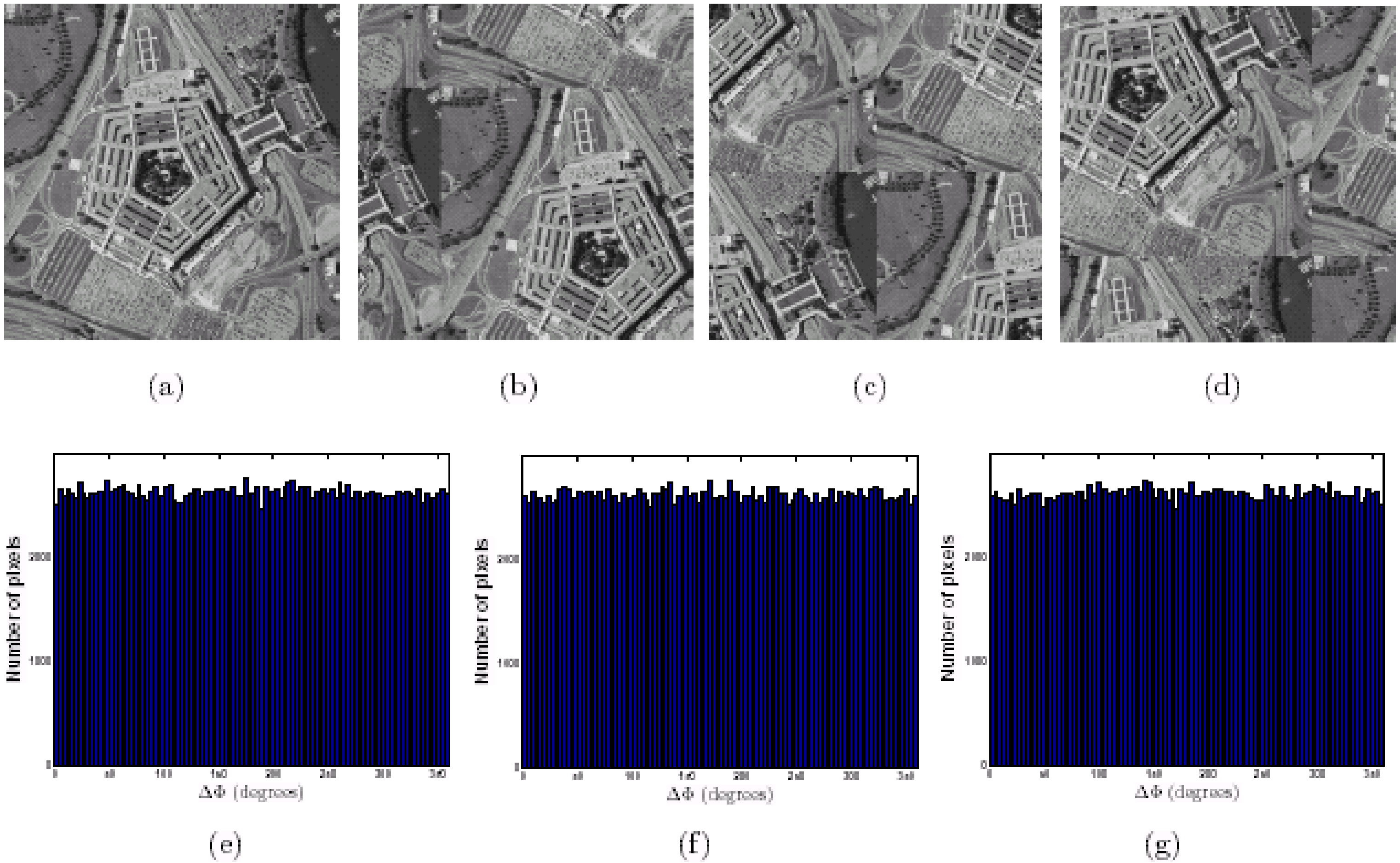
1. $0 \leqslant |\mathrm{NGC}(\mathbf{u})| \leqslant 1$
2. Invariance in affine changes in illumination

- Nominator

1. R_i : reward pixel locations with strong edge response
2. $\Delta\Phi(\mathbf{u},\mathbf{x}) = \Phi_1(\mathbf{x}) - \Phi_2(\mathbf{x}+\mathbf{u})$: Dirac-like shape, ability to reject outliers

ROBUST FFT-BASED IMAGE REGISTRATION

NORMALIZED GRADIENT CORRELATION   A CLOSER LOOK (2)



FFT-BASED SCALE INVARIANT IMAGE REGISTRATION

PRINCIPLES

- Assume two images related by a translation, rotation and scaling

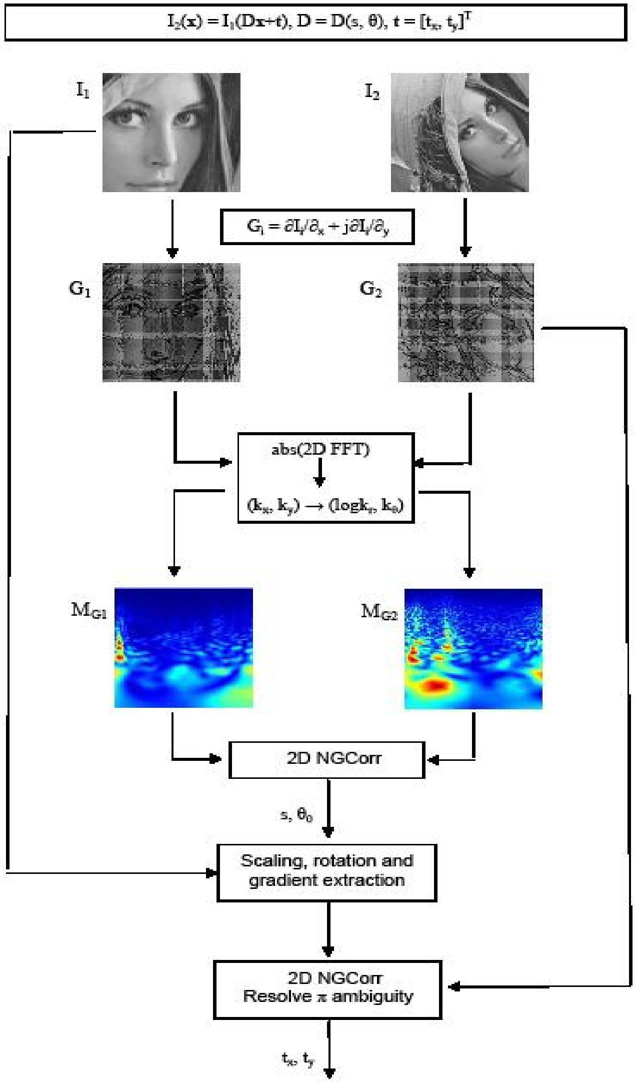
$$I_2(\mathbf{x}) = I_1(D\mathbf{x} + \mathbf{t}) \qquad D = s\Theta \text{ and } \Theta = \begin{bmatrix} \cos\theta_0 & \sin\theta_0 \\ -\sin\theta_0 & \cos\theta_0 \end{bmatrix}$$

- Translations do not affect the FFT magnitudes M_i. Resampling on a log-polar grid reduces scaling and rotation to a 2D translation which can be estimated using correlation

$$M_2(\mathbf{k}_l) = M_1(\mathbf{k}_l - [\log s, \theta_0]^T)$$

- Compensate for scaling and rotation and estimate the residual translation using 2D correlation in the spatial domain.

ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION



OUR METHOD

1. We replace I_i with G_i, compute the 2D FFT and keep its magnitude $M_{G_i} = |\widehat{G}_i|$.
   - M_{G_i} captures the frequency response of image salient structures solely. Areas of constant intensity levels induce low frequency components which hinder the estimation process.
   - No need for image windowing to reduce the effect of boundaries whose registration corresponds to zero motion.
   - Using filters with band-pass selection properties to compute G_i, reduces the effect of high-frequency noise and aliasing.
2. We replace standard correlation with Normalized Gradient Correlation.

ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION

RESULTS

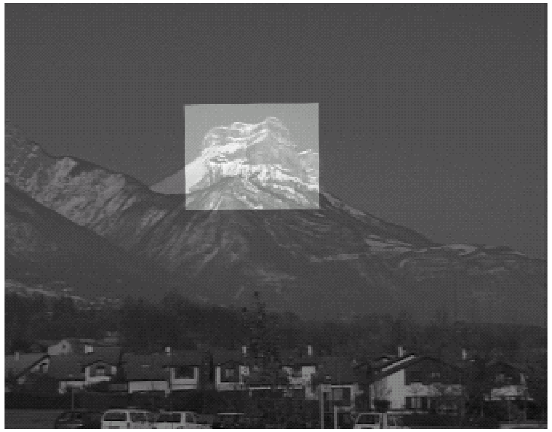
- Common belief: FFT-based schemes unable to cope with complex scenes
- State-of-the-art: Scale factors up to 2-2.5
- Our scheme: Scale factors up to 6

EXAMPLE 1

$(s, \theta_0) = (5.89, 33.2°),\ (\hat{s}, \hat{\theta}_0) = (5.85, 31.6°)$

ROBUST FFT-BASED SCALE-INVARIANT IMAGE REGISTRATION

EXAMPLE 2 $(s, \theta_0) = (3.97, 0.0°), \ (\hat{s}, \hat{\theta}_0) = (4.01, 0.7°)$



EXAMPLE 3 $(s, \theta_0) = (4.36, 46.0°), \ (\hat{s}, \hat{\theta}_0) = (4.26, 45.7°)$

Thanks for your attention !