Image Registration With Fourier-Based Image Correlation: A Comprehensive Review of Developments and Applications

12 minutes presentation with focus on mathematics and algorithmic aspects.

- We need to define a tool to produce the slides (overleaf/canva/powerpoint) and create the template (this second part can be done later)

- The goals of the survey, and how the authors have organized their article toward those.
  - What are we gonna talk about here?
  - Which parts of the paper talks about it?
  - How much time should we spend on this topic?
  - Are we going to do it all together or are we defining a responsible for developing the text and slides? Who?
  - Input the presentation text here?

- Summarize what the surveyed method is about, what key processing steps are involved with a mention on the mathematical tools and concepts used in the different steps, and give some examples of challenges at the various processing steps.
  - What are we gonna talk about here?
  - Which parts of the paper talks about it?
  - How much time should we spend on this topic?
  - Are we going to do it all together or are we defining somebody responsible for developing the text and slides? Who?
  - Input the presentation text here?

- Choose one application mentioned in the survey, pick 2 (or more) cited papers and summarize their results (metrics, cohort,...)
  - Which ones are we gonna talk about here?
  - Which parts of the paper talks about it?
  - How much time should we spend on this topic?
  - Are we going to do it all together or are we defining somebody responsible for developing the text and slides? Who?
  - Input the presentation text here?

- Your personal insights on the surveyed method: would you feel at ease implementing it? Would you recommend trying it on other data or applications ? How would you improve it?
  - What are we gonna talk about here?
  - How much time should we spend on this topic?
  - Are we going to do it all together or are we defining somebody responsible for developing the text and slides? Who?
  - Input the presentation text here?

Vitinho: Introduction and one paper
Benattinho: Methodology (including a bit of subpixel)
Julianzinho: the other paper, ease of implementing and other possible applications
Quem ler é brocha
Vc sabia que era isso e ainda fez mó trampo pra ler

- First Slide

Good morning everyone, I'm Vitor and our group is composed by Rafael, Julian and me and we'll present the paper "Image Registration With Fourier-Based Image Correlation: A comprehensive review of developments and applications".

  - Summary

I'll make an introduction about the article, Rafael will talk about the background necessary to understand the theoretical part and also give a brief presentation about subpixels methods, Julian and I, we will show two applications and finally Julian will make the conclusions.

- Introduction

The goal of this paper was to create a detailed review of the literature related to Fourier-based image correlation given an overview of the fundamentals, developments and some applications of image registration with Fourier-based image.

  - What's image registration?

So first things first, what's image registration?

Image registration is the process of overlaying images of the same scene, discovering matching spots and spatially align them to minimize the desired error.

Generally this field is divided into two categories the feature-based methods and the area-based methods:

- The feature-based methods extract features like edges, regions, shapes, corners and match them using similarity measures or geometric attributes
- On the other hand, the area-based methods use the intensity information to match areas or regions.

  - What's Fourier-based image correlation?

Fourier-based image correlation is an area-based method that aligns images using the information in the frequency domain.

This method is known for your theoretical accuracy, insensitivity to the frequency dependent noise and intensity contrast and high computational efficiency.

Since the use of fast Fourier Transform and the Phase Correlation technique for image registration, Fourier-based image correlation has played an important role in various applications and also has experienced rapid development with a large number of refinements being proposed.

- Methodology (including a bit of subpixel)

The Fourier-based methods can be divided into two main types of correlation forms, the **cross-correlation in frequency domain** and the **phase correlation.**

The cross-correlation form uses the correlation theorem, which states that the correlation between two signals can be calculated through the inverse fourier transform of the product of the fourier transform of one signal by the complex conjugate of the other signal. The use of fourier transform avoids iterative searching. The cross-correlation is bigger in the points where the convolution of the signal is bigger, as so, we would have a peak in the point of displacement of the image.
.However, this operation can result in an incorrect dominant peak if the images have different magnitudes.

What really makes sense is to find phase correlation between these two signals, where we would do the same operation, but dividing the product of the fourier transformed signals by the magnitude of this product as we can see in the phase correlation equation.
The advantage of this operation is that we don't do any kind of normalization between the two images to avoid an incorrect peak, because now, we aren't taking into account the magnitude of the signals. There are studies in the phase correlation calculation to reduce the computational costs of it, using tools such as discrete cosine transform and transformation of the image to a uni-dimensional signal.

NEXT SLIDE

These two methods only give us the translation between images. In a similar approach, assuming s as an isotropic scaling factor and theta-zero we can write the magnitude of the fourier transform of G as a function of these factors and the magnitude of F. Using the log-polar representation on both sides we get to a function where we can easily estimate the value of theta-zero and s.

NEXT SLIDE

The two translation estimation methods presented before are useful for integer pixel translations, which is not always the case. When the shift between two images is noninteger, the energy of the correlation peak spreads across the adjacent pixels. In order to estimate subpixel translation there are some calculation methods in the spatial domain and in the frequency domain. The methods in the spatial domain use the inverse fourier transform of the cross-power spectrum in the estimation of the peak while the ones in the frequency domain use the cross-power spectrum.

NEXT SLIDE

The peak centroid method calculates the centroid around the neighborhood of the integer-valued peak location, which causes the estimation to be biased toward integer values.

The interpolation methods use an interpolation function to estimate an offgrid maximum. There are many different interpolation functions used in the literature.

The upsampling methods upsample the images before calculating the cross-power spectrum, which is computationally expensive. The upsampling is done in the frequency domain by zero-padding the fourier transform of the images or by multiplying the image matrix by a larger upsampling matrix in the frequency domain. The matrix multiplication is less computationally expensive than the zero-padding method.

It's also possible to estimate a subpixel translation by minimizing the value of the gradient of the correlation function, that is, the inverse of the cross-power spectrum and find its zeros using the optimization approach.

NEXT SLIDE

The subpixel calculation in the frequency domain avoids the interference of the inverse fourier-transform reducing the interference of high-frequency components such as aliasing and noise.

The plane fitting method relies on the property that the phase of the cross-power spectrum in the continuous case represents a 2-D plane defined by the translation parameters. As so, the plane fitting methods tries to find plane that fits the best for the discrete case

The line fitting method separates each element of the cross-power spectrum, then we use a singular value decomposition which makes possible the recovery of the singular vectors of the left and right matrices by two independent lines.

NEXT SLIDE

The sawtooth estimation uses the property that the phase difference in the discrete case assumes a sawtooth signal shape based on the shift parameters. In practice, it's not a perfect sawtooth function, so it becomes a regression problem generally solved with least squares.

The frequency estimation is based on the unification of near frequencies in the cross-power spectrum using a multiple signal classifier algorithm, which is a bit computationally expensive.

In the optimization approach we minimize the cost function defined by the distance between the measured cross-power function and the theoretical one.
W represents here a frequency weighting or masking matrix.

NEXT SLIDE

Here we can see an absolute error comparison between some of these methods or interpolation functions. In general we see great results for the methods, the error stays generally below 0.3 pixels. The fourier-based methods had great results showing small mean absolute errors, although they had some outliers that are not much far away either.

The PEF here is the interpolation using the sinc derivation in 1-D channel, it has showed the best result between the spatial domain methods.

NEXT SLIDE (Application Examples)

JULIAN POCZONA NER KKKKKKKKKKKKKKKKKKKKKKKKKKKK

One example of the application of these tools is in the field of biometric authentication. In this paper by these people (K. Ito, H. Nakajima, K. Kobayashi, and T. Higuchi), "A Fingerprint Matching Algorithm Using Phase-Only Correlation".

In this work, a method of fingerprint authentication is developed, which improves the existing ones up to that date, especially in cases where the image is degraded, or the fingerprint is not of "good quality".

NEXT SLIDE

This method implements PC (here called POC - pass correlation only) to authenticate fingerprint matching. For this, they introduce Band-Limited POC, where they basically eliminate the high frequency phase components of the fingerprint, as these do not provide useful information.

NEXT SLIDE

In this slide, we can see how this new metric improve the peak detection. In the image above, we see the correlation field for two different images of the same fingerprint, where we clearly observe a peak. In the one below, we see the correlation for two different fingerprints, and we can see that in the case of POC there is also a peak, which can lead to false authentications.

NEXT SLIDE

This method was tested against a method that directly uses POC and a minutiae-based method (landmarks). We can expect that the landmarks are more ambiguous, this is why the method will be better, the method with POC contemplates all the texture information as opposed to just a few points like the other one.

Applications and conclusion.

These methods are simple to understand, but usually the difficulty comes from adapting them to a given problem. As the paper resumes, much progress has been made in recent years. However, there is still room for improvement in topics as:

- High-accuracy and high-efficiency subpixel methods.
- Effective task-oriented matching frameworks for different applications.
- Quantitive comparison in algorithms, implementations, and progressing frameworks

The solutions or improvements of these aspects propose a complicated challenge, they require a deep and exhaustive research work, so it is complicated for us to come up with possible lines of research or improvements.

However, what we did, is think of some possible improvements to the applied methods we presented. For example, in the method for cell tracking, they use the peak centroid method to find the movement of cells with subpixel precision, however, we have already seen that the PEF method has a better performance, so implementing this method would mean an improvement.

The field of applications of these methods is very wide, so we can imagine very diverse applications, such as in the area of autonomous vehicles, these vehicles use a lot of different sensors to be able to perceive the environment and from this make decisions. If we think about cameras, they will capture the scene from different angles, so one possible application, if not already implemented, could be to align these images in real time, to create a more complete and reliable detection of the scene. Adding lidar sensors, which are already implemented in autonomous vehicles, a 3D map of the scene can be created.

The paper also presents works where these methods are used as a tool to implement video stabilizers. If it’s not already implemented,in medicine this could be very helpfull in remote surgery, or laparoscopy, for example.