Advances in Knowledge-Based Technologies

Proceedings of the
Master and PhD Seminar
Summer term 2011, part 1

Softwarepark Hagenberg
SCCH, Room 0/2
15 April 2011

Software Competence Center Hagenberg
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 800
Fax +43 7236 3343 888
www.scch.at

Fuzzy Logic Laboratorium Linz
Softwarepark 21
A-4232 Hagenberg
Tel. +43 7236 3343 431
Fax +43 7236 3343 434
www.flll.jku.at

Program

Session 2. Chair: Roland Richter

10:15	Jean-Luc Bouchot: Some ideas for image registration based on a discrepancy minimization
10:45	Wolfgang Heidl: Active Learning from Soft Oracles

Technical Report
FLLL–TR–1102

FUZZY.LOGIC.LAB.LINZ

Some Ideas for Image Registration based on the Discrepancy Norm

15/04/2011

Jean-Luc Bouchot
Fuzzy Logic Laboratorium Linz
email: jean-luc.bouchot@jku.at

Abstract — In this talk we empirically show the feasibility of the discrepancy norm used in the context of image registration. We base our approach on a log-polar image representation which allows to convert any scale or rotational transformations into shift. A two steps optimization procedure which combines global optimization, based on the Direct algorithm, with local optimization techniques is then done which should gives us a reasonable coarse registration. A perfect registration, allowing more complex transformation can then easily be carried over based on the fast and efficient guess of our procedure.

Key words — *Discrepancy norm, image registration, optimization*

1 Introduction

Image registration is of main interest in the computer vision community. It is the problem of aligning two images coming from 1) different view points, 2) different sensors (known as multi-modal image registration), 3) different illumination, or different time of the day.

Nowadays, and since the development of the feature detection and matching, most of the methods try to connect keypoints from each of the input image. Those keypoints are of different kind (binary or floating points values), different size (128, 64 up to 512 dimensions) and allow different invariants (scale, rotation, translation, etc...) We refer the reader to [BETG08, CLSF10, DT05, HS88, Low04, Low99, MS05] for some examples of keypoints detection and keypoints description. The matching process, which goes out of the scope of this paper, is then done with some classical classifier (Tree based, Support Vector Machines, k-Nearest Neighbors for instance) after an outlier detection.

On the other hand, it is reasonable to think that having more information, or more control points, should allow to improve the robustness of the registration. This is the idea we want to develop here: image registration can be done directly on the pixel intensities, without the need of keypoints. Some efforts have been done in the past year for this task and we refer the reader to [BBL02] for some example. Of particular interest to us is the work of Wolberg and Zokai [WZ00, ZW05] and their log-polar registration. It indeeds allow to transfer zooming and rotation into shiftings. This idea combined with the monotonicity property of the discrepancy norm allows a general framework for the tasks of image registration.

This paper is structured as follows. We first give some basics about log-polar image representation and its property. Then we review the definition and property of the discrepancy norm as introduced in [NW87] and extended by Moser [Mos09] later. Sec. 3 will deal with some use-cases and examples and finally conclusion and future work will be given in Sec. 4.

2 Background

Before we study deeper the use of the discrepancy norm for image registration, we give some motivations for our work, starting with some basics on the log-polar transform (Sec. 2.1) and then motivate the choice of the discrepancy norm for this task, based on the work of Moser [Mos09] (Sec. 2.2).

2.1 Log-Polar transform

The log-polar transform was first used in Wolbergs and Zokais work[WZ00]. It converts a 2 dimensional cartesian image into an image with polar coordinate system. The log-polar transform needs a starting point to fix the origin of the polar coordinate system. Denote by c this point whose cartesian coordinates are (x_c, y_c) then the log polar transform is defined as:

$$LPT_c[I](\rho, \theta) = I(x_c + e^\rho \cos\theta, y_c + e^\rho \sin\theta) \tag{1}$$

where I is the image being analysed.

The term log-polar comes from the fact that the change of variable $\rho = \log r$ yields a classical polar representation. An example of application is shown on Figs. 1 and 2. Eq. 1 is of particular interest in the context of image registration as it allows to transform scaling and orientation changes into shift in this new space. Denotes Z_s a zooming function and R_α a rotation, we have:

$$Z_s[I](x,y) = I(e^s x, e^s y) \tag{2}$$

$$R_\alpha[I](x,y) = I\left(\begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\right) \tag{3}$$

If we assume now that an image is being rotated and zoomed, we write

$$\widetilde{I} = Z_s\left[R_\alpha[I]\right] \tag{4}$$

and we can compute the LPT of this image which yields to:

$$LPT[\widetilde{I}](\rho,\theta) = LPT[I](\rho + s, \theta + \alpha) \tag{5}$$

Eq. 5 shows that rotation and scale changes in the spatial cartesian domain are converted in shifts in the log-polar domain. This property is going to be usefull to us to develop our algorithm. But before giving the different steps, we first want to introduce a new (dis-)similarity measure for image registration.

2.2 Discrepancy Norm

While the discrepancy measurements has been studied for a very long time and dates back to Hermann Weyl's theory [Wey16], its interest for pattern recognition [NW87] and vision community is relatively recent. Proofs and more details on this topic can be found in [Mos09], and we will here recall only the properties and definitions which are interesting for our task.

The computation of the discrepancy norm on a 2 dimensional image I can be done with the use of integral images with the following formula:

$$\|I\|_D = \max_{\delta_1,\delta_2 \in \{-1,1\}} \max_{m,n \in \mathbb{Z}} \left| \sum_{i \in \mathcal{I}_{\delta_1}(m)} \sum_{j \in \mathcal{I}_{\delta_2}(n)} I_{i,j} \right| \tag{6}$$

where we have $\mathcal{I}_{\delta_k}(l) = \{i : \delta_k i \leq \delta_k l\}$

The autocorrealation function $\Delta[I](\mathbf{t},\lambda) = \|I \circ T_{\lambda\mathbf{t}} - I\|_D$, where $T_{\lambda\mathbf{t}}$ represents a translation of parameter $\lambda\mathbf{t}$, $\mathbf{t}$ being a vector, has an interesing monotonicity property for positive signals:

$$\forall I \geq 0,\ \forall \lambda_1, \lambda_2 \in \mathbb{R},\ |\lambda_1| \leq |\lambda_2| \Rightarrow \Delta[I](\mathbf{t},\lambda_1) \leq \Delta[I](\mathbf{t},\lambda_2) \tag{7}$$

This property is illustrated on Fig 3. One clearly sees the monotonic behaviour of the autocorrelation function. The different colors show two different size of the cropped area. This property 7 is going to be the key of our algorithm presented in the next section.

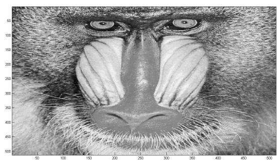

Figure 1: Image of a friendly monkey

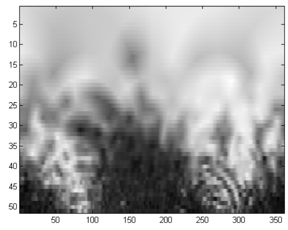

Figure 2: Log-polar transformation of the center

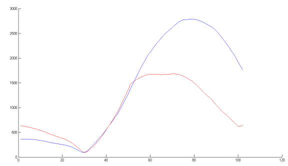

Figure 3: Monotonicity behavior in the scale direction

3 Contents of the talk

Some tests have been done on two different tasks:

- Localization. This is the problem of finding an image in another with different scale and rotation. Typical applications are in the field of orthogonal images or any aerial image.
- Alignment. This is the problem of aligning two images of the same object but taken from different conditions. This would be the case in some medical applications where the images could come from MRI and SPECT for instance. This allows a user to combine information contained in both images into a single representation.

Examples of both use cases will be given in the talk with more details on the procedure.

4 Conclusion - Future

This talk describes a first appraoch to image registration using the discrepancy norm. However some efforts have to be done in order to improve the whole process:

- How do we choose the center of the log polar registration? It seems clear out of the experiments, that this parameter has a critical impact on the accuracy and that only small displacements will yield completely different transformations. This idea chosen so far is the use of a sliding window, which is time consuming.
- How can we get a more uniform grid? At the moment, there is a critical bias between the points at the center which have a high importance and the one far away from it. This implies a loss of information with the radius of transformation getting bigger.
- How can we improve the optimization procedure to be perfectly adapted to the discrepancy norm?

A last point that is worth mentioning is that the discrepancy norm could also be used to find the ideal center of the transformation. However, this is the chicken and egg problem: knwoing the center allows us to find the scale and rotation of the images, and knowing those allows us to align (in translation) the corrected images.

References

[BBL02] A. Bodnarova, M. Bennamoun, and S. Latham. Optimal gabor filters for textile flaw detection. *Pattern Recognition*, 35:2973–2991, 2002.

[BETG08] Herbert Bay, Andreas Ess, Tinne Tuytelaars, and Luc Van Gool. Surf: Speeded up robust features. *Computer Vision and Image Understanding*, 110:346 – 359, 2008.

[CLSF10] M. Calonder, V. Lepetit, C. Strecha, and P. Fua. Brief: Binary robust independent elementary features. In *11th European Conference on Computer Vision (ECCV)*, 2010.

[DT05] Navneet Dalal and Bill Triggs. Histograms of oriented gradients for human detection. In Cordelia Schmid, Stefano Soatto, and Carlo Tomasi, editors, *IEEE Comp. Soc. Conf. on Computer Vision and Pattern Recognition*, volume 2, pages 886 – 893, June 2005.

[HS88] C. Harris and M. Stephens. A combined corner and edge detector. In *Proc. of the 4th Alvey Vision Conf.*, pages 147 – 151, 1988.

[Low99] David G. Lowe. Object recognition from local scale-invariant features. In *Proc. of the Int. Conf. on Computer Vision*, volume 2, pages 1150 – 1157, Washington, DC, USA, 1999. IEEE Computer Society.

[Low04] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60:91 – 110, November 2004.

[Mos09] B. Moser. Similarity measure for image and volumetric data based on Hermann Weyl's discrepancy measure. *IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI)*, 99:DOI 10.1109/TPAMI.2009.50, 2009.

[MS05] Krystian Mikolajczyk and Cordelia Schmid. A performance evaluation of local descriptors. *IEEE Trans. on Patt. Anal. Mach. Intell.*, 27(10):1615 – 1630, 2005.

[NW87] H. Neunzert and B. Wetton. Pattern recognition using measure space metrics. Technical Report 28, Universität Kaiserslautern, Fachbereich Mathematik, November 1987.

[Wey16] H. Weyl. Über die Gleichverteilung von Zahlen mod. Eins. *Math. Ann.*, 77:313–352, 1916.

[WZ00] G. Wolberg and S. Zokai. Robust image registration using log-polar transform. In *Image Processing, 2000. Proceedings. 2000 International Conference on*, pages 493 –496 vol.1, 2000.

[ZW05] S. Zokai and G. Wolberg. Image registration using log-polar mappings for recovery of large-scale similarity and projective transformations. *Image Processing, IEEE Transactions on*, 14(10):1422 –1434, oct. 2005.

Active Learning from Soft Oracles

Wolfgang Heidl WOLFGANG.HEIDL@PROFACTOR.AT
Profactor Research, 4407 Steyr-Gleink, AUSTRIA

1. Introduction

The idea behind Active Learning (AL) is to induce more accurate models using less training samples by letting the learning algorithm chose instances to be annotated in a supervised learning setting. Interest in AL has been driven by the fact that while training instances as such are often abundant, acquiring labels for those instances is costly.

There exist a number of different strategy to identify the most informative sample for querying, including *uncertainty sampling* (Lewis & Catlett, 1994), *query-by-committee* (Abe & Mamitsuka, 1998), *expected model change* and *expected error reduction* (Settles, 2009). All of these strategies are selecting samples solely from the viewpoint of the learning algorithm. The effectiveness of those sampling is demonstrated in simulation studies, where only an initial set of prelabeled data is provided to start training and labels for further instances are selectively queried. Thereby, it is implicitly assumed that the oracle always provides the correct answer and that there is no influence from the query strategy to the oracle (Donmez & Carbonell, 2008).

However, if we think of a realistic case, where the task is to learn from humans neither of the two assumptions holds. Recent results from an empirical study (Baldridge & Palmer, 2009) in fact suggest that effectiveness of AL is dependent on oracle expertise and that for a non-expert random sampling is superior to uncertainty sampling. The authors conclude that dealing with variations in annotators may be more important than devising better selection strategies.

2. Our Approach

To indicate the difference to the predominant assumption of machine-like oracles where AL is only concerned with selecting the most informative sample we propose to consider annotators are *Soft Oracles* whose performance is influenced by the selection strategy. Specifically we assume that Soft Oracles are

- Fallible: they will not always answer correctly, their answers are increasingly random for difficult annotations.
- Learning: during the annotation process oracles become acquainted with the task at hand. Their efficiency and accuracy is to some extent dependent on reassuring, easy samples.

We present an approach that avoids asking too difficult questions that will result in almost random answers and hinder 'coming into the game' of oracles. Instead of choosing between sampling the most uncertain sample and random sample selection (Baldridge & Palmer, 2009) we propose a method that naturally varies between the two in a way adaptive to the performance of soft oracles. This way, not only can we deal with different levels of initial expertise, the sampling strategy is continuously adapted as oracles learn during an annotation session.

The problem of noisy oracles has already been recognized as an important issue in real-world applications of AL, (Sheng et al., 2008), (Dekel & Shamir, 2009). The predominant approach in that area is to consider multi-oracle settings, where individual oracle fidelity can be gauged against others and annotation load allocated to optimize overall annotation cost. Instead of focusing on avoiding bad oracles we aim at optimizing utilization of a given *soft* oracle. Better utilization of oracles can of course also be applied in multi-oracle setting and combined with cost-optimal allocation.

Like uncertainty sampling and query by committee, our query strategy relies on a measure of distance to the decision boundary, which for the former is given by the classification (un)certainty and for the letter by the amount of (dis)agreement among committee members. We assume that those measures of distance produced by the classifier are related to difficulty for an oracle to label an instance and that the relation can be esti-

This work has been supported by the FEMtech program of The Federal Ministry for Transport, Innovation and Technology under FFG grant No. 318113. It reflects only the authors' views.

mated from the labeled portion of the data. Instead of selecting the sample closest to the decision boundary, our strategy is to sample according to a predefined level/distribution of subjective oracle difficulty.

3. Results

We apply our proposed active learning strategy to sample selection in a visual inspection experiment. Previously, in a similar experiment (Heidl et al., 2011) sample selection was guided by a fixed decision model and the relevant sample area around the decisions boundary estimated from pilot experiment data. Now active learning is employed to select samples form a large pool, starting from a preselected set of 10% of queried samples.

To reach a fair comparison of model accuracy between fixed and active sample selection we estimate the expected prediction error on an independent set of 10000 samples from the same distribution. Since no true label is known for those data we use the approach recently proposed by Domez et al. (Donmez et al., 2010), which estimates the prediction error of unlabled samples from the overlap of a bimodal Gaussian distribution fitted to the discrimant value distributions of the samples. Using our active learning approach, the expected error on an independent set of data is almost halved from 17.6% to 9.4%. The resulting effect sizes of gender differences in induced decision trees are between 7% and 35% larger than those previously reported, with the largest effect size of $d = 1.02$ for tree entropy.

References

Abe, N., & Mamitsuka, H. (1998). Query learning strategies using boosting and bagging. *Proceedings of ICML'98: Fifteenth International Conference on Machine Learning*. Madison, WI, USA: Morgan Kaufmann Publishers Inc.

Baldridge, J., & Palmer, A. (2009). How well does active learning actually work?: Time-based evaluation of cost-reduction strategies for language documentation. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing* (pp. 296–305). Stroudsburg, PA, USA: Association for Computational Linguistics.

Dekel, O., & Shamir, O. (2009). Good learners for evil teachers. *Proceedings of the 26th Annual International Conference on Machine Learning* (pp. 233–240). New York, NY, USA: ACM.

Donmez, P., & Carbonell, J. G. (2008). Proactive learning: cost-sensitive active learning with multiple imperfect oracles. *Proceeding of the 17th ACM conference on Information and knowledge management* (pp. 619–628). New York, NY, USA: ACM.

Donmez, P., Lebanon, G., & Balasubramanian, K. (2010). Unsupervised supervised learning i: Estimating classification and regression errors without labels. *J. Mach. Learn. Res.*, *11*, 1323–1351.

Heidl, W., Thumfart, S., Lughofer, E., Eitzinger, C., & Klement, E. P. (2011). Decision tree-based analysis suggests structural gender differences in visual inspection. *Proceedings of AIA2011, IASTED International Conference on Artificial Intelligence and Applications* (pp. 142–149). Innsbruck, Austria.

Lewis, D. D., & Catlett, J. (1994). Heterogeneous uncertainty sampling for supervised learning. *In Proceedings of ICML'94: Eleventh International Conference on Machine Learning* (pp. 148–156). Morgan Kaufmann.

Settles, B. (2009). *Active learning literature survey*Computer Sciences Technical Report 1648). University of Wisconsin–Madison.

Sheng, V. S., Provost, F., & Ipeirotis, P. G. (2008). Get another label? improving data quality and data mining using multiple, noisy labelers. *Proceeding of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 614–622). New York, NY, USA: ACM.