Finger-vein Sample Compression in Presence of Pre-Compressed Gallery Data

Tamara Lipowski, Babak Maser, Jutta Hämmerle-Uhl, and Andreas Uhl
Department of Computer Sciences
University of Salzburg, AUSTRIA
uhl@cosy.sbg.ac.at

***Abstract*—Compression settings for sample (probe) finger vein data in case of already pre-compressed gallery data are investigated. Inhomogeneous compression scenarios are assessed where probe data can be compressed with different compression technique and compression ratio compared to gallery data using 4 lossy compression schemes, 2 finger vein recognition schemes, and 2 data sets. Results obtained indicate that in case of JPEG2000 pre-compressed gallery, also sample images should be compressed in the same manner, while for JPEG pre-compressed gallery, the optimal sample compression setting depends on the dataset, on the target compression ratio, and on the recognition scheme employed.**

***Index Terms*—finger vein recognition, lossy compression, pre-compressed gallery, mixed compression schemes.**

I. Introduction

Since biometric data has begun to be stored digitally, minimising file size has been an important concern, in order to increase both storage and transmission efficiency. Therefore, it is highly profitable to look into the best suited compression methodology for a given biometric modality, application scenario or even dataset, and to look further into possible interference between compression technology and feature extraction / template generation algorithms.

The certainly most relevant standard for compressing image data relevant in biometric systems is the ISO/IEC 19794 standard suite on Biometric Data Interchange Formats where in the most recently published version (ISO/IEC 19794-9:2011 for vascular data), JPEG, JPEG_LS, and JPEG2000 are included for lossy compression (see clause 8.3.13).

The ANSI/NIST-ITL 1-2011 standard on "Data Format for the Interchange of Fingerprint, Facial & Other Biometric Information" (former ANSI/NIST-ITL 1-2007) only supports JPEG2000 for applications tolerating lossy compression. Of course, to achieve a significant data reduction, only lossy compression techniques are admissible.

Apart from standardisation, a variety of independent studies dealing with compression and the respective impact on biometric recognition performance exist for various modalities (see e.g. [1] for corresponding references). So far, compression impact on vein-based biometric recognition schemes has hardly been investigated. JPEG compression of hand vein data was considered in the context of a general investigation of hand vein recognition algorithms' robustness [2]. The closest work to this current manuscript is [1], in which three ISO/IEC still image compression standards have been applied to a single public finger vein data set and eventual interference of a collection of feature extraction / template generation schemes with different types of compression artifacts have been studied in detail. Results have indicated that the optimal compression scheme depends on the target compression ratio as well as on the employed recognition algorithms.

In this paper, we systematically investigate a more advanced topic in the context of finger vein data compression. When studying literature on compressing biometric data, we typically are able to identify two considered scenarios, i.e. either compressing probe and gallery (i) with the same compression technique to the same compression ratio or (ii) leaving one of the two uncompressed (e.g. the gallery) and compressing the probe with different techniques and compression ratios. We term the first scenario as "homogeneous compression" as identical techniques and compression ratios are applied to both probe and gallery. These two scenarios do not necessarily correspond to a real world setting. In this work, we assume that the gallery data has been already stored in some specific compressed form after enrollment (specific with respect to compression technique and compression ratio). We need to decide (maybe at some point later in time where compression technology has advanced compared to the time of enrollment) in which way the probe data should be compressed to result in the best recognition accuracy. This question is also of general (i.e. non-biometric) interest as it is unclear if the presence of compression artefacts of different or similar nature (as generated by different or identical compression techniques) play a decisive role for recognition accuracy in this context. As in this scenario neither compression technique nor compression ratio of probe and gallery data need to match we term this scenario as "inhomogeneous compression". Up to our knowledge, this has neither been investigated in any work on biometric data compression nor in other pattern recognition related work.

Section 2 introduces the still image compression standards as considered in this work and briefly reviews the finger vein recognition algorithms as used in the paper. In Section 3 we first explain the experimental settings (including the two finger vein datasets used in experimentation) and subsequently discuss the results of a large corpus of experiments. Section 4 presents the conclusions of this paper.

II. Image Compression and Finger Vein Recognition

We consider four image compression techniques with increasing compression ratios (i.e. ratio between original file size and file size after compression, shown on x-axis of plots in the experimental section) up to 110 – to cover also samples with significantly reduced size as eventually required for wireless transmission – using the respective default configurations unless stated otherwise:

JPEG (JPG): The well-known (ISO/IEC IS 10918-1) DCT-based method. By adjusting the divisors in the quantisation phase, different compression ratios can be achieved. We adjust the quality parameter iteratively to achieve a file size closest to the desired compression ratio. The Matlab implementation is used.

JPEG 2000 (J2K): This wavelet-based standard (ISO/IEC IS 15444-1) is also a part of the DICOM standard where it replaced lossless JPEG compression. Results typically do not generate block-based artifacts as the original DCT-based JPG standard. J2K facilitates explicit rate control, i.e. target bitrates are met with high accuracy. We use JJ2000 version 5.1 available at https://code.google.com/p/jj2000/. For J2K, we additionally employ three variants of region of interest coding (ROI, i.e. the pixels corresponding to the finger): First, the classical variant where all the ROI data is coded into the bitstream before the background data ("J2K (ROI)"); second, where also resolution Level 0 of the background is encoded together with the ROI data ("J2K (SL0)"); and third, where the first 4 resolution levels of the background is encoded with the ROI data ("J2K (SL4)").

JPEG-XR (JXR): Is based on Microsoft's HD Photo and is known to produce higher quality than JPEG, but provides faster compression than JPEG 2000. In the default configuration the Photo Overlay/Overlap Transformation is only applied to high pass coefficients prior to the Photo Core Transformation (ISO/IEC IS 29199-2). We adjust quantisation levels iteratively to achieve a target bitrate closest to the desired one. Software available at https://jxrlib.codeplex.com/ is used in experiments.

BPG: The "Better Portable Graphics" algorithm is based on a subset of the H.265 (HEVC, ISO/IEC 23008-2) video compression standard. We adjust quantisation levels iteratively to achieve a target bitrate closest to the desired one. The employed software can be downloaded from https://bellard.org/bpg/.

Preprocessing is conducted in accordance to [3]. For *finger alignment*, we simply mask out background pixels (setting them to 0). This is followed by a normalisation step, i.e. rotation compensation. For *image enhancement*, we apply CLAHE as the final stage of High Frequency Emphasis Filtering (HFE).

To foster reproducible research, we have only included *feature extraction* techniques in this study for which basic implementations are publicly available.

Maximum Curvature (MC [4]) aims to emphasise only the centre lines of the veins and is therefore insensitive to varying vein width. The first step is the extraction of the centre positions of the veins. Therefore the local maximum curvature in the cross-sectional profiles, based on the first and second derivatives, are determined. Afterwards each profile is classified as being concave or convex where only local maxima in concave profiles indicate valid centre positions of the veins. Then a score according to the width and curvature of the vein region is assigned to each centre position, which is recorded in a matrix called locus space. Due to noise or other distortions some pixels may not have been classified correctly at the first step, thus the centre positions of the veins are connected using a filtering operation. Finally binarisation is done by thresholding using the median of the locus space.

For matching the binary feature images we adopted the approach in [3]. As the input images are not registered to each other and only coarsely aligned (rotation is compensated), the correlation between the input image and in x- and y-direction shifted versions of the reference image is calculated. The maximum of these correlation values is normalised and then used as final matching score.

In contrast to MC, the **SIFT** key-point [5] based technique uses information from the most discriminative points as well as considering the neighbourhood and context information of these points by extracting key-points and assigning a descriptor to each key-point. We employ an additional key-point filtering as described in [3], where keypoints close to the finger / background boundary are discarded in addition to eventual background keypoints.

SIFT matching is done using the keypoint descriptors – the keypoint with the smallest distance to the reference keypoint is the matched one if the distance is below a threshold, otherwise there is no match. To resolve the problem with ambiguous matches (i.e. one keypoint may have small distances to more than one other point) the classical ratio threshold scheme is used: A match is only valid if the distance of the best point match is at least k (threshold) times smaller than to all other points.

For MC the software of B.T. Ton [1] is used, while for SIFT, feature extraction and matching software is used as provided by VL_Feat SIFT[2].

III. Experiments

A. Experimental Settings

For experimental evaluations in this paper, we use the following two publicly available finger-vein datasets, both containing data compressed in lossless manner:

UTFVP: University of Twente Finger Vascular Pattern Database (UTFVP) [6], consisting of a total of 1440 images, taken from 60 subjects, 6 fingers per subject and 4 images per finger. The images have a resolution of 672×380 pixels

[1]Publicly available on MATLAB Central: http://www.mathworks.nl/matlabcentral/fileexchange/authors/57311

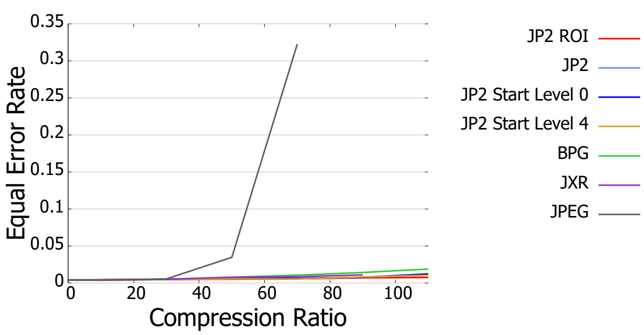
[2]http://www.vlfeat.org/

with 8 bit grey scale depth.

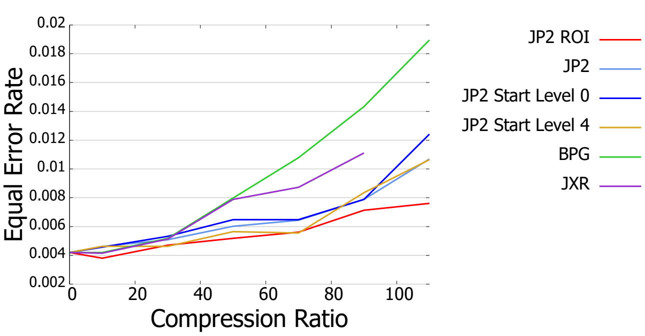
SDUMLA-HMT: This multi-modal dataset was collected at Shandong University, Jinan, China. 106 subjects, including 61 males and 45 females with age between 17 and 31, provided traits for face, finger vein, gait, iris and fingerprints [7]. SDUMLA-HMT is available at http://mla.sdu.edu.cn/sdumla-hmt.html. The finger-vein dataset consists of 6 fingers per subject, 6 images per finger, 3816 images, 320×240 pixels with 8 bit grey level.

The test procedure of the FVC2004 was adopted to determine the EER (shown on y-axis in plots). While we have also computed ZeroFNMR values and generated corresponding plots, space limitations prevent us from showing those. However, these exhibit the same trends as seen in the EER data.

For the inhomogeneous compression scenario, we consider gallery data to be compressed to compression ratios 10 and 30 (for good and medium quality) using JPEG and JPEG2000, respectively. The choice of these two compression schemes is motivated by their inclusion in the ISO/IEC 19794-6 standard as the "old" (JPEG) and more recently suggested compression scheme. Thus, gallery data compressed in this manner (e.g. JPEG for datasets acquired some years ago, JPEG2000 for datasets acquired more recently) can be considered a realistic scenario.



(a) with JPEG



(b) without JPEG

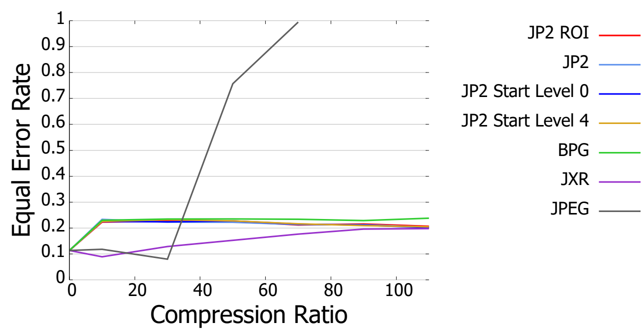
Fig. 1: MC on UTFVP data (EER), JPEG 2000 gallery compression (ratio 10).

B. Experimental Results

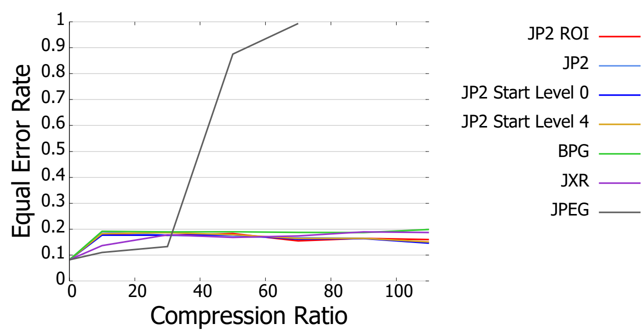
Fig. 1(a) represents the typical behaviour seen for both data sets, compressing the gallery data with JPEG2000 ratio 10 or 30 and JPEG with ratio 10 and using MC recognition, respectively. JPEG (the purple line) is only to keep up until compression ratio 30 and exhibits exploding error rates subsequently. Results of this type (although in the homogeneous scenario) have been observed also in [1] – the strong block-based distortions as generated by JPEG at low bitrates mislead the feature extraction algorithms to detect vein structures at the block borders causing high matching errors.

When discarding the JPEG results (Fig. 1(b)), we observe that recognition accuracy drops for increasing the compression ratio of the probe image in monotonous fashion, typically exhibiting the best behaviour for some JPEG2000 variant, followed by JXR and BPG giving the worst results in terms of EER.

However, there are some exceptions as shown in Figs. 2(a) and (b), in particular when the gallery data is compressed with JPEG to compression ratios 10 and 30.



(a) gallery ratio 10



(b) gallery ratio 30

Fig. 2: SIFT with increasing on UTFVP data (EER), JPEG gallery compression.

What we observe in some rare cases (typically seen for SIFT recognition), is that JPEG probe compression gives the best recognition results (up to ratio 30 as shown in Fig. 2). This is due to the match of the strong artefacts, which are of course identical both of JPEG type in probe and gallery data.

In the following tables (Tables I - II) we display the best compression scheme at a specified probe compression ratio

TABLE I: Lowest EER for Gallery JPEG Compression

UTFVP: Gallery JPEG Compression Rate 10						
	10	30	50	70	90	110
MC	J2K	J2K (SL4)	J2K (ROI)	J2K (SL4)	J2K (SL0)	J2K (ROI)
Diff	0.00 (JXR)	0.00 (J2K ROI)	0.00 (J2K SL4)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00(J2K)
SIFT	JPEG	JPEG	J2K	J2K (ROI)	J2K (SL4)	J2K (SL0)
Diff	0.10 (J2K ROI)	0.14 (J2K)	0.00 (J2K ROI)	0.00 (J2K)	0.00 (J2K SL0)	0.00 (J2K)
SDUMLA: Gallery JPEG Compression Rate 10						
	10	30	50	70	90	110
MC	J2K	J2K (SL4)	J2K (SL4)	J2K	J2K (ROI)	J2K (SL0)
Diff	0.00 (J2K SL4)	0.00 (J2K)	0.00 (J2K)	0.00 (J2K SL4)	0.00 (J2K SL0)	0.00 (J2K ROI)
SIFT	JPEG	J2K	J2K (ROI)	J2K (ROI)	J2K (SL0)	J2K (ROI)
Diff	0.00 (J2K)	0.00 (J2K SL0)	0.00 (J2K SL0)	0.00 (J2K SL0)	0.00 (J2K ROI)	0.00 (J2K)
UTFVP: Gallery JPEG Compression Rate 30						
	10	30	50	70	90	110
MC	J2K	J2K (ROI)	J2K (ROI)	J2K	J2K (ROI)	J2K (ROI)
Diff	0.00 (J2K SL4)	0.00 (JPEG)	0.00 (J2K)	0.00 (J2K SL4)	0.00 (J2K)	0.00 (J2K)
SIFT	JPEG	JPEG	J2K	J2K (ROI)	J2K (SL0)	J2K (SL0)
Diff	0.03 (BPG)	0.04 (J2K ROI)	0.01 (J2K SL0)	0.01 (J2K SL0)	0.00 (J2K SL4)	0.00 (J2K)
SDUMLA: Gallery JPEG Compression Rate 30						
	10	30	50	70	90	110
MC	JPEG	J2K (SL4)	J2K (SL4)	J2K (SK4)	J2K (SL4)	J2K
Diff	0.00 (JXR)	0.00 (J2K)	0.00 (JXR)	0.00 (J2K)	0.00 (J2K)	0.00 (J2K SL4)
SIFT	JPEG	JXR	JXR	J2K (ROI)	J2K	J2K (SL0)
Diff	0.00 (JXR)	0.00 (J2K SL4)	0.00 (J2K SL0)	0.00 (JXR)	0.00 (J2K ROI)	0.00 (J2K ROI)

(i.e. 10 - 110) in terms of EER, given the type and compression ratio of the gallery data (i.e. JPEG or JPEG2000 and compression ratio 10 and 30) for both data sets. Additionally, in the line below each entry, we list the difference in EER to the second-best technique (rounded to two positions after the decimal point) and the second-best technique itself in brackets.

There are very clear trends visible in the data. In the case of MC recognition, there are only two cases in which a non-JPEG2000 variant represents the best probe compression option (i.e. gallery JPEG compression with ratio 30 for SDUMLA data for probe compression ratio 10 (JPEG is best) and gallery JPEG2000 compression with ratio 30 for UTFVP data for probe compression ratio 10 (again JPEG is best)). For SIFT recognition, the optimal probe compression scheme is different from JPEG2000 more often. In particular, for SIFT, especially at lower ratios and JPEG gallery compression, we typically see non-JPEG2000 techniques (i.e. some DCT related scheme) being the best ones. While this seems to be surprising at first sight, there is an obvious explanation for that. For higher probe bitrates, DCT-type artefacts (i.e. as seen for JPEG, JXR, and BPG) are similar to those of the gallery images (which are compessed by JPEG), resulting in the effect that these schemes are found the be the best option. In case the probes' bitrate is further reduced, DCT-related artefacts get stronger than those in the gallery images and the better JPEG2000 quality results in better matching behaviour.

There is another very clear trend: In case of JPEG2000 gallery compression (Table II, in most cases some variant of JPEG2000 probe compression is the optimal choice (except for 3 cases which only occur at gallery compression with ratio 30). Furthermore, a clear majority of best JPEG2000 variants take advantage of region of interest coding.

It is also obvious, that these general observations do apply for both datasets, while result details (e.g. the actual JPEG2000 variant being best) are highly dependent on the considered dataset. Overall, the differences to the second-best ranked techniques are small. Significant differences in terms of EER are only seen for gallery compression with JPEG, SIFT-based recogntion and the UTFVP dataset. For all other settings, first and second ranked techniques exhibit fairly similar EERs.

Summarising it gets clear that inhomogeneous compression is more beneficial in case of gallery compression with JPEG and applied SIFT recognition. In this case a wide variety of (DCT-related) compression schemes represent the best solution for probe data compression at lower compression ratios, while JPEG2000 is the variant of choice for higher ratios. In case of JPEG2000 gallery compression, the inhomogeneous approach does not offer big advantages, only in few cases compression schemes different than JPEG2000 represent the best probe compression option. Furthermore it turns out that region of interest compression in JPEG2000 is beneficial in many settings.

IV. Conclusion

The experimental compression evaluations conducted in this paper indicate that it is beneficial under certain circumstances to use different compression techniques and compression ratios for probe and gallery images ("inhomogeneous compression"). These circumstances do follow a clear and deterministic pattern: For JPEG gallery compression and SIFT recognition, at least at the lower probe compression ratios compression schemes based on some DCT variant are the optimal choice. SIFT recognition explicitly looks at the neighbourhood of keypoints, thus similar compression artefacts are important for successful matching, while for MC the concentration is on the binarisation process which does not seem to be that compression artefact dependent. However, for higher compression ratios, the differences between probe and gallery images

TABLE II: Lowest EER for Gallery JPEG 2000 Compression

UTFVP: Gallery JPEG 2000 Compression Ratio 10						
	10	30	50	70	90	110
MC	J2K (ROI)	J2K (SL4)	J2K (ROI)	J2K (SL4)	J2K (ROI)	J2K (ROI)
Diff	0.00 (JXR)	0.00 (J2K ROI)	0.00 (J2K SL4)	0.00 (J2K ROI)	0.00 (J2K)	0.00 (J2K SL4)
SIFT	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)
Diff	0.00 (J2K ROI)	0.00 (J2K SL4)	0.00 (J2K ROI)	0.00 (J2K)	0.00 (J2K SL4)	0.00 (J2K)
SDUMLA: Gallery JPEG 2000 Compression Ratio 10						
	10	30	50	70	90	110
MC	J2K	J2K (SL4)	J2K (SL4)	J2K (SL4)	J2K (SL0)	J2K (ROI)
Diff	0.00 (J2K SL4)	0.00 (J2K)	0.00 (J2K)	0.00 (J2K SL0)	0.00 (J2K ROI)	0.00 (J2K SL0)
SIFT	J2K	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)
Diff	0.00 (JXR)	0.00 (J2K SL4)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00 (J2K ROI)
UTFVP: Gallery JPEG 2000 Compression Ratio 30						
	10	30	50	70	90	110
MC	JPEG	J2K (SL0)	J2K (ROI)	J2K (SL4)	J2K (ROI)	J2K (ROI)
Diff	0.00 (J2K)	0.00 (J2K ROI)	0.00 (J2K)	0.00 (J2K)	0.00 (J2K SL0)	0.00 (J2K SL4)
SIFT	J2K (ROI)	J2K (ROI)	J2K (ROI)	J2K (ROI)	JXR	J2K (SL4)
Diff	0.00 (J2K SL0)	0.00 (J2K SL4)	0.00 (J2K SL4)	0.00 (J2K SL0)	0.00 (J2K SL4)	0.00 (J2K ROI)
SDUMLA Gallery JPEG 2000 Compression Ratio 30						
	10	30	50	70	90	110
MC	J2K (ROI)	J2K	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (ROI)
Diff	0.02 (J2K)	0.01 (J2K SL0)	0.00 (J2K SL0)	0.00 (J2K SL0)	0.00 (J2K SL0)	0.00 (J2K SL4)
SIFT	JPEG	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)	J2K (SL0)
Diff	0.00 (J2K SL0)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00 (J2K ROI)	0.00 (J2K ROI)

get too large under compression with weaker compression schemes for SIFT to be successful any longer, and thus JPEG2000, the overall best technique, is then also the best option for JPEG compressed gallery data.

References

[1] V. Ablinger, C. Zenz, J. Hämmerle-Uhl, and A. Uhl, "Compression standards in fingervein recognition," in *Proceedings of the 9th IAPR/IEEE International Conference on Biometrics (ICB'16)*, 2016, pp. 1–7.

[2] C. Kauba and A. Uhl, "Robustness evaluation of hand vein recognition systems," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'15)*, Darmstadt, Germany, 2015, pp. 1–8.

[3] C. Kauba, J. Reissig, and A. Uhl, "Pre-processing cascades and fusion in finger vein recognition," in *Proceedings of the International Conference of the Biometrics Special Interest Group (BIOSIG'14)*, Darmstadt, Germany, Sep. 2014.

[4] N. Miura, A. Nagasaka, and T. Miyatake, "Extraction of finger-vein patterns using maximum curvature points in image profiles," *IEICE transactions on information and systems*, vol. 90, no. 8, pp. 1185–1194, 2007.

[5] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision (CVPR'99)*, vol. 2. IEEE, 1999, pp. 1150 – 1157.

[6] B. Ton and R. Veldhuis, "A high quality finger vascular pattern dataset collected using a custom designed capturing device," in *International Conference on Biometrics, ICB 2013*. IEEE, 2013. [Online]. Available: http://doc.utwente.nl/87790/

[7] Y. Yin, L. Liu, and X. Sun, "SDUMLA-HMT: A Multimodal Biometric Database," in *The 6th Chinese Conference on Biometric Recognition (CCBR 2011)*, ser. Springer Lecture Notes on Computer Science, vol. 7098, 2011, pp. 260–268.