The Complete Rank Transform:
A Tool for Accurate and Morphologically Invariant Matching of Structures

Oliver Demetz
demetz@mia.uni-saarland.de

David Hafner
hafner@mia.uni-saarland.de

Joachim Weickert
weickert@mia.uni-saarland.de

Mathematical Image Analysis Group
Saarland University
Saarbrücken, DE

Incorporating invariances is a popular tool to make computer vision systems more robust under real-world conditions. Since being invariant means ignoring something, every invariance has to come at the price of a loss of information. In our paper, we focus on severe illumination changes and descriptors that are invariant under such challenging conditions. The goal of our paper is to develop a novel descriptor that discards as little information as possible and carries the *maximally possible* amount of local image information under this invariance.

Intensity Order-based Transforms

Before describing our novel patch-based descriptor, let us briefly revisit two closely related descriptors [2, 3] that form the basis of our proposal.

The *rank transform* (RT) encodes for each pixel the position of its grey value in the ranking of all grey values in its neighbourhood. Practically, this rank is determined by counting the number of neighbours with a smaller grey value than the reference pixel, cf. Figure 1(b).

The *census transform* (CT) can be seen as an extension of the rank transform: Besides encoding the rank, it adds a spatial component by expressing the relationship between the central pixel and each of its neighbours explicitly. In practice, one bit is stored for each pixel of the neighbourhood: If the neighbour is smaller than the reference pixel, the bit is 1, otherwise it is 0, cf. Fig. 1(c). Finally, all bits are concatenated in one signature.

We propose a novel transform that extracts as much local information as possible from the input image data while preserving the same desired invariance against monotonically increasing illumination changes. To this end, we make use of the following observation: Not only the rank of the central intensity value, but even the *complete intensity order* of the considered patch is morphologically invariant.

Hence, we extend the rank transform in such a way that it incorporates this novel idea of encoding the *complete* intensity order. The construction of our novel *complete rank transform* (CRT) is straightforward: First, we compute the rank of each pixel in the considered neighbourhood patch by determining the number of pixels with smaller intensity. Next, we concatenate these ranks as for the census transform to obtain the complete rank descriptor. Mathematically, for a patch size k, the complete rank transform is the mapping $\boldsymbol{s}_{\text{CRT}} : \Omega \to \Pi_k \subset \{0,\ldots,k-1\}^k$, where the co-domain Π_k is the set of all possible rankings of k elements including multiple occurrences of the same intensity value. The cardinality of this set is the k-th ordered Bell number ($OBN(k)$).

4	14	83
4	25	88
3	15	65

(a) Intensity values.

	5	

(b) Rank.

1	1	0
1		0
1	1	0

(c) Census.

1	3	7
1	5	8
0	4	6

(d) Complete rank.

Figure 1: Illustration of the presented intensity order transforms (*(b)–(d)*) with a 3×3 neighbourhood patch (*(a)*), where the reference pixel is marked in grey.

Table 1: Comparison of the proposed intensity order transforms. The number of pixels in the considered neighbourhood is given by k.

transform	range D of one digit	signature length κ	spatial information	size of descriptor space
RT	$\{0,\ldots,k-1\}$	1	–	k
CT	$\{0,1\}$	$k-1$	✓	2^{k-1}
CRT	$\{0,\ldots,k-1\}$	k	✓	OBN(k)

Variational Optic Flow Model

As a proof of concept, we embed the novel signature in a generic $TV-L^1$-type variational model for optic flow [1] which assumes the signatures of corresponding pixels to coincide. We obtain the functional

$$E(u,v) = \int_\Omega \Big(\Psi\Big(\tfrac{1}{\kappa}|\boldsymbol{s}(\boldsymbol{x}+\boldsymbol{w}) - \boldsymbol{s}(\boldsymbol{x})|^2\Big) + \alpha\cdot\Psi\Big(|\boldsymbol{\nabla} u|^2 + |\boldsymbol{\nabla} v|^2\Big)\Big)\,\mathrm{d}\boldsymbol{x}, \tag{1}$$

whose minimiser is the sought flow field $(u,v)^\top : \Omega \to \mathbb{R}^2$. The first term (*data term*) penalises differences between the vector-valued signature $\boldsymbol{s} : \Omega\times[0,\infty) \to D^\kappa$ at position $\boldsymbol{x} := (x,y,t)^\top$ in the first frame and its corresponding one at $\boldsymbol{x}+\boldsymbol{w} := (x+u, y+v, t+1)^\top$ in the second frame. The second term (*smoothness term*) penalises the magnitude of the gradient $\boldsymbol{\nabla} := (\partial_x, \partial_y)^\top$ of the flow field.

Results

Our novel CRT consistently outperforms the existing signature types. Figure 2 shows its unconditional invariance against γ-changes. On the Middlebury benchmark, our method ranked 43.5th, and on the KITTI Vision Benchmark it ranked 9th at time of submission.

[1] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert. High accuracy optical flow estimation based on a theory for warping. In *Proc. ECCV*. 2004.

[2] M. Pietikäinen, A. Hadid, G. Zhao, and T. Ahonen. *Computer Vision Using Local Binary Patterns*. Springer, 2011.

[3] R. Zabih and J. Woodfill. Non-parametric local transforms for computing visual correspondence. In *Proc. ECCV*. 1994.

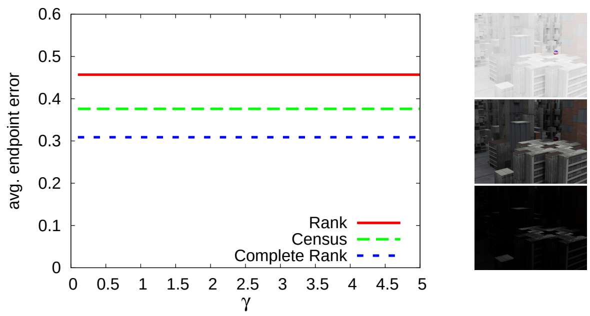

Figure 2: Behaviour under γ-changes. The plots show the results of our method under γ variations of the second frame of the *Urban2* sequence ($\gamma = \{0.1, 1, 3\}$).