Danafar et al. *EURASIP Journal on Advances in Signal Processing* 2012, **2012**:124
http://asp.eurasipjournals.com/content/2012/1/124

EURASIP Journal on Advances in Signal Processing
a SpringerOpen Journal

CORRECTION Open Access

Correction: novel kernel-based recognizers of human actions

Somayeh Danafar, Alessandro Giusti* and Jürgen Schmidhuber

In our paper *Novel Kernel-based Recognizers of Human Actions* [1], several sentences should be corrected as indicated below. The corrections do not affect the basic results of the paper.

- Section 2 (Related works), Subsection 2.1 (Features for Action Recognition). The last three sentences of the second paragraph should read as follows:

  Jhuang et al. [20] present a hierarchical supervised method with spatio-temporal, gradient and flow filters organized in various layers of complexity. In the last layer a multi-class SVM recognizes the human action.

- Section 2 (Related works), Subsection 2.1 (Features for action recognition). The last paragraph should read as follows:

  The approach is taken further by Schindler and Van Gool [24], who investigated the detection of actions from very short sequences called snippets. Two separate pathways for motion and shape are considered. Motion is modeled by means of optical flow, computed for different directions and scales. Shape is represented by Gabor filter responses. MAX-pooling and comparison with a set of templates (learned using PCA) yield high-level feature vectors, which are classified through SVMs. In our approach, we feed our classification algorithm by such a powerful feature descriptor, independently computed for each pair of frames.

- Section 2 (Related works), Subsection 2.2. (Classification for action recognition). The second paragraph should read as follows:

  Previous work [25] proposes 2D spatio-temporal compound features that are learned in a weakly-supervised approach using a data mining algorithm. Several researchers have explored unsupervised methods for motion analysis. Hoey [26] uses a multilevel dynamic Bayesian Network as an unsupervised classifier of facial expressions. Zhong et. al. [27] propose an unsupervised method to detect unusual activities in videos, by comparison with action prototypes. An alternative approach [28] detects unusual activity by spectral clustering and a hierarchical observation Hidden Markov Model. Boiman and Irani [29] explain a video sequence using patches from a database; as dense sampling of the patches is necessary in their approach, the resulting algorithm is very time-consuming and unpractical for action recognition. Wang et al. [30] adopt spectral clustering to cluster a large set of human action images. In this context, shape features are used to compute distances, by means of a linear programming approach. Niebels et al. [11] deal with unsupervised learning of human action categories. Each action is represented by a probability distribution on spatio-temporal features. Action classes are modeled by latent topic models, such as probabilistic Latent Semantic Analysis and Latent Dirichlet Allocation.

- The following sentence should be appended to the caption of figure 3 (page 7).

  Illustration adapted from reference [24], ©2008 IEEE. Reprinted, with permission, from Proc. CVPR 2008.

- Section 3 (The Maximum Mean Discrepancy). The second sentence should read as follows.

  Recent work [5,6,8] studies the embedding of random variables into a Reproducing Kernel Hilbert Space (RKHS), by using kernels that take into account information from higher order statistics.

- Section 3 (The Maximum Mean Discrepancy). Definition 1 should read as follows.

*Correspondence: alessandrog@idsia.ch

Let $P \in \mathcal{P}$ be a Borel probability measure which is defined on a separable metric space $\mathcal{X}$. A positive definite and bounded kernel k denoted as $k : \mathcal{X} \times \mathcal{X} \rightarrow \mathbb{R}$ can be used to map $\mathcal{P}$ to an RKHS $\mathcal{F}$. For each $x \in \mathcal{X}$ and continuous function $\phi(x)$, a kernel is an inner product computation, $k(x, x') :=< \phi(x), \phi(x') >_{\mathcal{F}}$. The P-based expectation of $\phi(x)$ is the so-called mean element μ_P [8,9]:

$$\begin{aligned} \mu_P : \mathcal{P} &\rightarrow \mathcal{F} \\ P &\mapsto \int_{\mathcal{X}} \Phi(x) dP. \end{aligned}$$

- Section 3. The second paragraph after definition 3 should read as follows.

  Theorem 2 defines the distance measure between probability distributions. However, we also need to measure whether the resulting distance is statistically significant for asymptotic distributions P and Q. In a two-sample test this is done by comparing the null hypothesis, $\mathcal{H}_0 : P = Q$, against the alternative hypothesis $\mathcal{H}_1 : P \neq Q$; a significance threshold is then obtained.

- After Theorem 6, the last sentence in the paragraph should read as follows.

  This quantile can be estimated by means of the bootstrap method [8,9].

Received: 21 May 2012 Accepted: 21 May 2012
Published: 21 June 2012

References

1. S Danafar, A Giusti, J Schmidhuber, Novel kernel-based recognizers of human actions. EURASIP J. Adv. Signal Process. — Special Issue Video Anal. Human Behavior Understanding. (2010). doi:10.1155/2010/202768

doi:10.1186/1687-6180-2012-124
Cite this article as: Danafar *et al.*: **Correction: novel kernel-based recognizers of human actions.** *EURASIP Journal on Advances in Signal Processing* 2012 **2012**:124.