Anisotropic Partial Differential Equation Based Video Saliency Detection

Wai Lam Hoo Chee Seng Chan
University of Malaya, Malaysia

UCF Sports Dataset

Introduction

There has been numerous saliency detection models for video saliency detection. At the same time, PDEs have been used successfully in many low-level image processing tasks such as image denoising, inpanting etc, and recently in more complex tasks such as image saliency detection. However, the work is not suitable for video saliency detection task as the original LESD model

(i) does not consider the orientation and motion information contain in the video; and
(ii) uses the center-prior.

This work proposes a novel method to generate the static saliency map based on the **adaptive nonlinear PDEs model** with refinements. Experiments on various human action datasets show that our proposed model performs favourably against the conventional methods.

Proposed method

Given a set of saliency seeds S and their corresponding score function $u(x; y; i) = u0; i\epsilon S$, we mathematically formulate saliency detection as an evolutional PDEs with initial condition such as:

$$\begin{aligned} \frac{\partial u_i}{\partial t} &= G(u, \nabla u, |\nabla u|), \text{on } I\times(0,\infty); \\ u(x,y,0) &= u_0 \text{ on } I; \\ \partial u_n &= 0, \text{on } I\times(0,\infty). \end{aligned}$$

However, it does not give a reliable information in the presence of flow-like structures. So, structure tensor is selected to rotate the flow towards the orientation of interesting features.

$$\begin{aligned} S_\sigma &= (\textstyle\sum_{i=1}^{n} \nabla \mathrm{u}_{i\sigma} \nabla \mathrm{u}_{i\sigma}^{T}) \\ &= \begin{bmatrix} \sum_{i=1}^{n} u_{ix\sigma}^2 & \sum_{i=1}^{n} u_{ix\sigma}\, u_{iy\sigma} \\ \sum_{i=1}^{n} u_{ix\sigma}\, u_{iy\sigma} & \sum_{i=1}^{n} u_{iy\sigma}^2 \end{bmatrix}. \end{aligned}$$

The description of the local gradient characteristics is improved with these new gradient features, and it's smoothed version of S_σ , can be represented as:

$$J_p = K_\rho * S_\sigma = \begin{bmatrix} j_{11} & j_{12} \\ j_{21} & j_{22} \end{bmatrix}.$$

In order to create a truly anisotropic scheme, the nonlinear diffusion tensor is used replacing the diffusivity function $G(\cdot)$ with the combination of two types of novel tensors as follows:

$$\frac{\partial u_i}{\partial t} = div((K_1(J_\rho) + \alpha K_2(J_\rho) \nabla \mathrm{u_i})$$

To incorporate the high-level prior into the diffusion process, another regularization term, $\eta(\cdot)$ is introduced such that:

$$\frac{\partial u_i}{\partial t} = div((K_1(J_\rho) + \alpha K_2(J_\rho) \nabla \mathrm{u_i}) + \eta(u(i) - d(i)), i\epsilon S.$$

The situation when the saliency evolution is stable is considered, therefore only find a solution to the following PDE:

$$\begin{aligned} \frac{\partial u_i}{\partial t} &= 0, \text{on } I\times(0,\infty); \\ u(x,y,0) &= u_0 \text{ on } I; \\ \partial u_n &= 0, \text{on } I\times(0,\infty), i\epsilon S. \end{aligned}$$

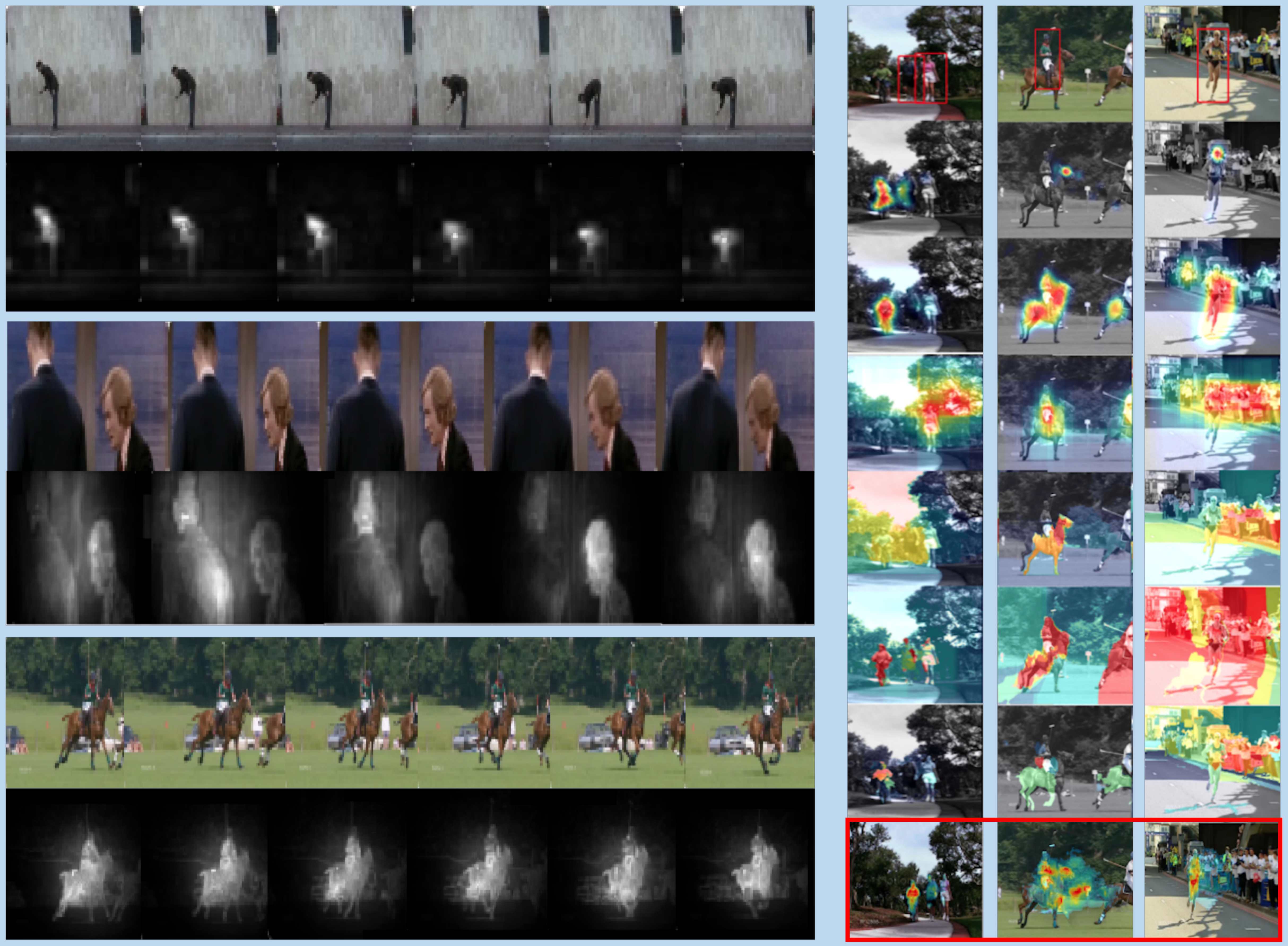
Therefore, given a video frame, the saliency detection task rounds to the problem of to achieve a stable state for visual attention diffusion:

$$u_i^{n+1} = u_i^n + \Delta t\, div((K_1(J_\rho) + \alpha(K_2(J_\rho) \nabla \mathrm{u}_i^n) + \eta(u(i) - d(i)), i\epsilon S$$

The final direction map is defined as $d(i) = L * C_{rg} * C_{by} * Te * u_f(i)$, where C_{rg} and C_{by} are color features, Te is texture feature and $u_f(i)$ is foreground score; and the final (video) saliency map, S is represented as:

$$\begin{aligned} S &= S_{un} \otimes S_{re}, \\ S_{un} &= \psi(S_s, S_{m_p}), \\ S_{re} &= \psi(S_{s_p}, S_m). \end{aligned}$$

Results



Results on UCF Sports dataset, Weizmann dataset and Hollywood dataset.

Conclusion

This paper proposes a novel video saliency detection method inspired by PDEs. Particularly, we introduce a novel method to generate static saliency map based on the adaptive nonlinear PDEs model. Experimental results had shown the effectiveness of the proposed method in three public human action datasets when compared to the conventional solutions.