Co-occurrence based Foreground Detection with Hypothesis on Degradation Modification in Severe Imaging Conditions

周　文俊†, 金子　俊一†, 佐藤　雄隆‡, 橋本　学⁂, 梁 棟⁑

Wenjun ZHOU†, Shun'ichi KANEKO†, Yutaka SATOH‡, Manabu HASHIMOTO⁂ and Dong LIANG⁑

†: 北海道大学

‡: 産業技術総合研究所

⁂: 中京大学

⁑ : Nanjing University of Aeronautics and Astronautics

zhouwenjun@ssc.ssi.ist.hokudai.ac.jp

1. Overview of CPB

As an extension of "pixel to pixel" structure, in CPB we design a "pixel to block" structure to reduce the time computing and the proposed CPB [1] includes two processes: training process and detecting process as shown in Fig.1.

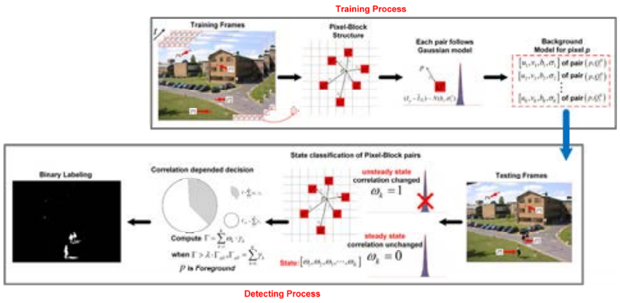

Fig.1. Overview of CPB

1.1. Training process

First, we select the blocks which are the highest K components of the Pearson's product-moment correlation coefficient $\gamma(p,Q^B)$ as the supporting blocks Q_k^B for the target pixel p. For the selected K pixel-block pairs, we build a statistical model using a single Gaussian distribution as defined in the following expression:

$$\Delta_k \sim N\left(b_k,\sigma_k^2\right) \quad \Delta_k = \bar{I}_p - \bar{I}_{Q_k} \quad , \qquad (1)$$

where $\bar{I}_p$ is the intensity of the pixel p at t frame and $\bar{I}_{Q_k}$ is the average intensity of the block Q_k^B at t frame.

1.2. Detecting process

For each pair $\left(p,Q_k^B\right)$, a binary function for identifying its steady or unsteady state can be defined as follows:

$$\omega_k = \begin{cases} 1 & if \left|\left(p-Q_k^B\right)-b_k\right| \ge \eta\cdot\sigma_k \\ 0 & otherwise \end{cases} \quad , \qquad (2)$$

where $\left|\left(p-Q_k^B\right)-b_k\right|$ represents a bias in the intensity difference between the real value and the modeled parameter b to estimate the steady or unsteady state of each pair $\left(p,Q_k^B\right)$. To further evaluate the state of pixel p, we introduce a weighting γ_k into the previous binary definition. We define the following evaluation criterion to classify the target pixel into the foreground class as: **if** $\Gamma \ge \lambda\cdot\Gamma_{all}$ **,** p is *foreground*; **else** p is *background*.

2. HoD modification

2.1. Hypothesis on Degradation

In practice, after a long-time utilization of initial CPB background model in an unlearned sequence, the expected relative relation of the pixel-block pair might be broken. In other words, initial CPB model might generate a degradation with the passage of time, then some "noise" might arise in detecting process. Here, we define such assumption as "Hypothesis on Degradation" and name the "noise" in detecting process as "hypothetical noise": (1) the hole surrounded by the detected foreground pixels, which is estimated as the background and we named it 'NaB'; (2) the dot surrounded by the non-detected pixels, which is estimated as the event and we named it 'NaE'. Fig.2 shows an example of the hypothetical noise.

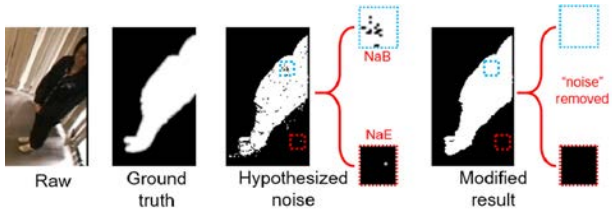

Fig.2. Description of hypothesized noise

2.2. Broken pixel-block pairs detection

Fig.3 describes an overview of the proposed HoD. First we need to detect the broken elemental pairs in pixel-block structure of the hypothetical noise. In this study, we assume that the larger γ could hold a higher weight in the trained pixel-block structure and such pair would be more likely to affect the state of pixels. We propose a weight-based decision rule to detect the wrong pair:

$$if \ \gamma_m \ge \bar{\gamma}, then \ \left(p,Q_m^B\right) \ is \ broken. \qquad (3)$$

Where $\left(p,Q_m^B\right)$ is the 'wrong' pair, which is unsteady state of NaE or steady state of NaB. In the case of NaE, it is defined by use of the total number of

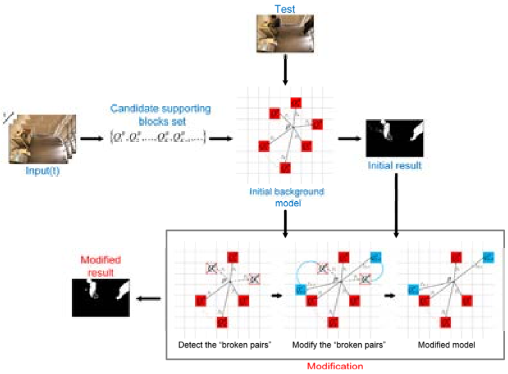

Fig.3. Overview of HoD modification

unsteady pairs $M = \sum_{k=1}^{K} \omega_k$ as follows:

$$\bar{\gamma} = \frac{1}{M} \sum_{k=1}^{K} \gamma_k \cdot \omega_k = \frac{1}{M} \Gamma. \tag{4}$$

In the other hand, for NaB, it is defined as follows:

$$\bar{\gamma} = \frac{1}{K-M} \sum_{k=1}^{K} \gamma_k \cdot (1-\omega_k) = \frac{1}{K-M} (\Gamma_{all} - \Gamma), \tag{5}$$

then we record these broken pairs for the next process.

2.3. Structure modification

Then, we try to exchange the broken pair by new one which is kept as a spare pair in the training process and remove the hypothesized noise.

3. Experiments

We compare the proposed CPB and CPB+HoD with six different foreground detection methods: GMM and KDE, which are two well-known traditional algorithms, and four state of the art techniques: IMBS [2], T2FMRF-UV [3], ViBe [4] and SuBSENSE [5], especially SuBSENSE is one of the top-ranked methods in CDW-2012 dataset at present and T2FMRF-UV is a foreground extraction algorithm specifically for dynamic backgrounds. Fig.4 shows examples of foreground detection in the challenges: illumination changes and background motion, respectively. Table1 and Table2 list the results of the performance measurements.

4. Conclusions

We developed a prospective background model with hypothesis on degradation modification (HoD) for foreground detection under severe imaging conditions. It was designed to handle the problem of strong background changes in reality. With the help of HoD, we further improve the robustness of CPB and stabilize the effectiveness in the long-term use.

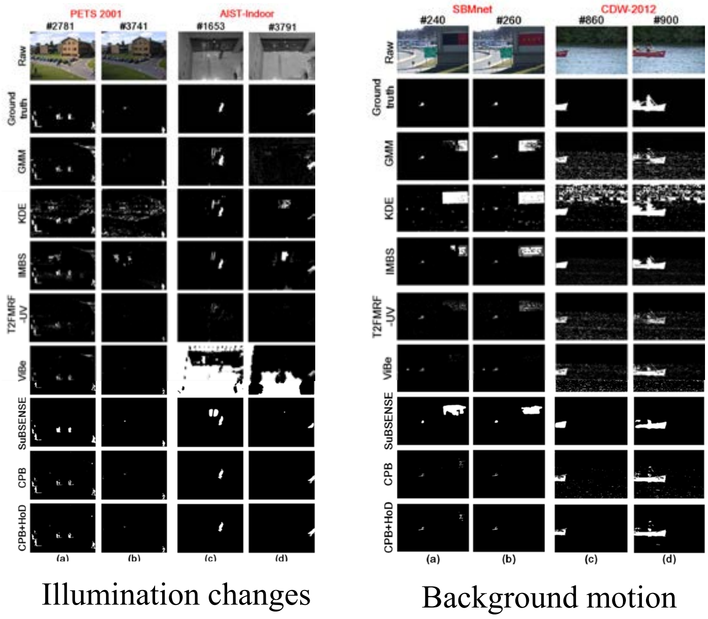

Fig.4. Foreground detection results in illumination changes and background motion, respectively.

Table1. Comparison in illumination changes

Datasets	PETS 2001				AIST			
Methods	Precision	Recall	F-measure	PSNR	Precision	Recall	F-measure	PSNR
GMM	0.6465	0.9508	0.7697	39.46	0.6523	0.9207	0.7636	40.57
KDE	0.5181	0.8836	0.6531	17.77	0.5896	0.6944	0.6377	38.16
IMBS	0.5162	0.8841	0.6518	16.20	0.5760	0.6923	0.6288	36.36
T2FMRF-UV	0.5818	0.8365	0.6863	34.94	0.6382	0.5818	0.6087	45.65
ViBe	0.7059	0.8821	0.7842	43.42	0.5005	0.5146	0.5074	9.11
SuBSENSE	0.9008	0.8840	0.8923	54.11	0.5864	0.7047	0.6401	37.14
CPB	0.9566	0.7517	0.8418	56.05	0.8651	0.8181	0.8409	53.14
CPB+HoD	0.9652	0.7562	0.8480	56.39	0.8668	0.8227	0.8442	53.31

* Note that red entries indicate the best in measurement, and blue entries indicate the second best.

Table2. Comparison in background motion

Datasets	SBMnet				CDW-2012			
Methods	Precision	Recall	F-measure	PSNR	Precision	Recall	F-measure	PSNR
GMM	0.5151	0.5196	0.5174	26.92	0.6748	0.7024	0.6883	21.71
KDE	0.4962	0.4856	0.4909	21.67	0.6584	0.8630	0.7468	17.22
IMBS	0.5095	0.5118	0.5107	30.09	0.7315	0.8911	0.8035	21.60
T2FMRF-UV	0.5508	0.5179	0.5338	35.38	0.6797	0.6114	0.6438	23.49
ViBe	0.6427	0.5368	0.5850	35.16	0.8114	0.7821	0.7965	28.02
SuBSENSE	0.5018	0.5033	0.5025	27.62	0.9766	0.7649	0.8573	30.80
CPB	0.7653	0.5118	0.6133	36.64	0.9283	0.7730	0.8436	32.61
CPB+HoD	0.7973	0.5214	0.6350	37.39	0.9809	0.7830	0.8708	34.16

* Note that red entries indicate the best in measurement, and blue entries indicate the second best.

References

[1] W. Zhou, S. Kaneko, D. Liang, M. Hashimoto, Y. Satoh, Background subtraction based on co-occurrence pixel block pairs for robust object detection in dynamic scenes, IIEEJ transactions on image electronics and visual computing 5 (2) (2017) 146–159.

[2] D. Bloisi, L. Iocchi. "Independent multimodal Background subtraction," CompIMGE (2012) 39-44.

[3] Z. Zhao, T. Bouwmans, X. Zhang, Y. Fang, A fuzzy background modeling approach for motion detection in dynamic backgrounds, Multimedia and signal processing 346 (2012) 177–185.

[4] O. Barnich, M. Van Droogenbroeck, Vibe: A universal background subtr-action algorithm for video sequences, IEEE Transactions on Image processing 20 (6) (2011) 1709–1724.

[5] P.L. St-Charles, G. A. Bilodeau, R. Bergevin. "Subsense: A universal change detection method with local adaptive sensitivity," IEEE Transati-ons on Image Processing 24 (1) (2015) 359-373.