Moving Shadow Tracking in VR Interaction

A novel optimized approach

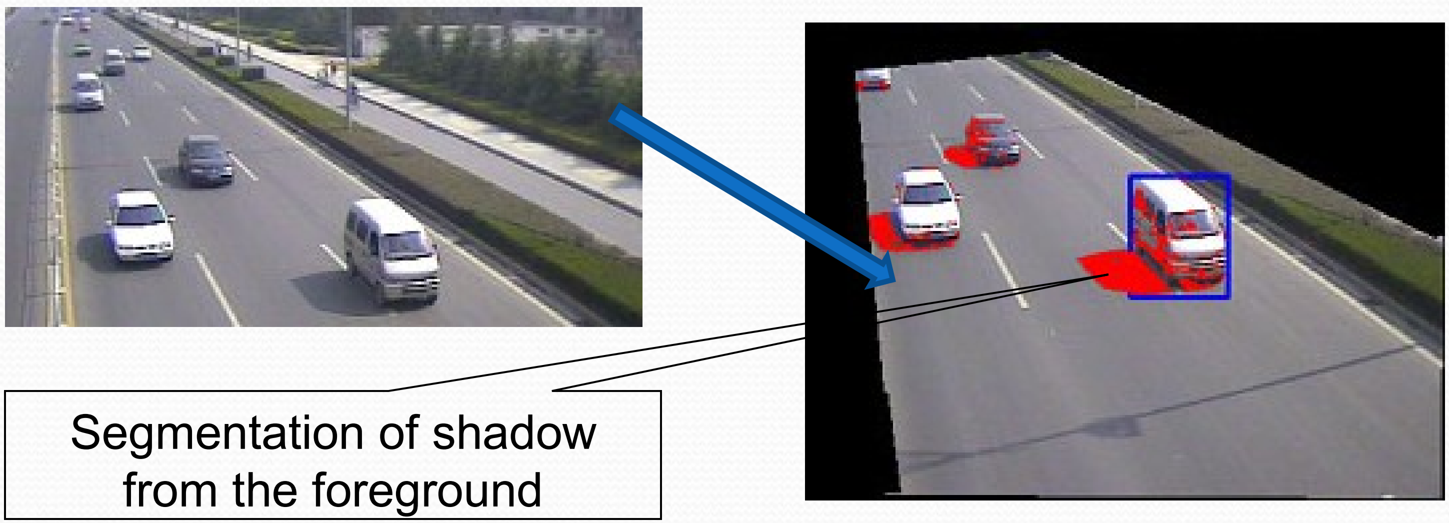
Haipeng Cai

Outline

- Moving Shadow Tracking - the generals
- The two-step discriminant
- Improve the classical GMM (Gaussian Mixture Model)
- MSTVRI - the whole flow
- Results
- Summary

MST - the generals

- MST in the video motion detection
  - Remove shadow so as to improve the quality of motion detection



  - MST by use of shadow’s chromatic feature is an effective way with low performance loss

MST - the generals

- MST in Video-based VR Interaction
  - The feature of application – we only care the shadow but not that which casts it
    - *The video frame does only contain the shadow, rather than the moving objects, mostly people who would interact with the video scene*
  - Based on the shadow's characteristic of motion, the shadow itself could be treated as special moving object as in the video motion detection [*Prati A. 2001*]

The two-step discriminant

- The proportion of gray between pixel in the background and that in the shadow area

  [Jehan-Besson 2001]

in the RGB color space

$$
\begin{aligned}
Gray_s &= \alpha R_s + \beta G_s + \gamma B_s \\
&= \alpha kR_b + \beta kG_b + \gamma kB_b \\
&= k(\alpha R_b + \beta G_b + \gamma B_b) = kGray_b
\end{aligned}
$$

- Step-I

$$
SP_1(x_t) = \begin{cases} 1, & (X_r < b_r \wedge X_g < b_g \wedge X_b < b_b) \\ & \wedge (T_s \le Gray(X_t) \le T_l) \\ 0, & otherwise \end{cases}
$$

The two-step discriminant

- As a background area is casted into shadow, the saturation will decrease appreciably with only very trivial change on the part of its Value in the HSV (Hue-Saturation-Value) color space [Prati A. 2003]

- Step-II

$$SP_2(x_t) = \begin{cases} 1, (\alpha \le \frac{I^V(xt)}{B^V(xt)} \le \beta) \wedge \\ \quad (I^S(xt) - B^S(xt) \le TS) \wedge \\ \quad (\left| I^H(xt) - B^H(xt) \right| \le TH) \\ 0, \quad otherwise \end{cases}$$

Improve the classical GMM

- The classical framework *[Grisman&Stauffer 2000]*
  - Background modeling – the Gaussian Mixture

$$P(X_t) = \sum_{i=1}^{K} \omega_{i,t} * \frac{1}{(2\pi)^{n/2} |\Sigma_{i,t}|^{1/2}} e^{-\frac{1}{2}(X_t - \mu_{i,t})^T (\Sigma_{i,t})^{-1} (X_t - \mu_{i,t})}$$

$$i = 1, 2, \ldots, k$$

  - Model Adaptation

$$\omega_{i,t} = (1-\alpha)\omega_{i,t-1} + \alpha M_{i,t}$$

$$\mu_{i,t} = (1-\rho)\mu_{i,t-1} + \rho x_t$$

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho (x_t - \mu_{i,t})^T (x_t - \mu_{i,t})$$

$$\rho = \alpha \eta (x_t | \mu_{i,t}, \sigma_{i,t})$$

  - The real background model filter

$$B = \arg\min_b \left(\sum_{k=1}^{b} W_k > T_{sw}\right)$$

Improve the classical GMM

- The idea – optimization by simplification
  - Cut off the variance in the model adaptation

$$\sigma_{i,t}^2 = (1-\rho)\sigma_{i,t-1}^2 + \rho(x_t-\mu_{i,t})^T(x_t-\mu_{i,t})$$

to be omitted

  - Remove the probability factor

Almost equivalence, esp. in terms of our specific application

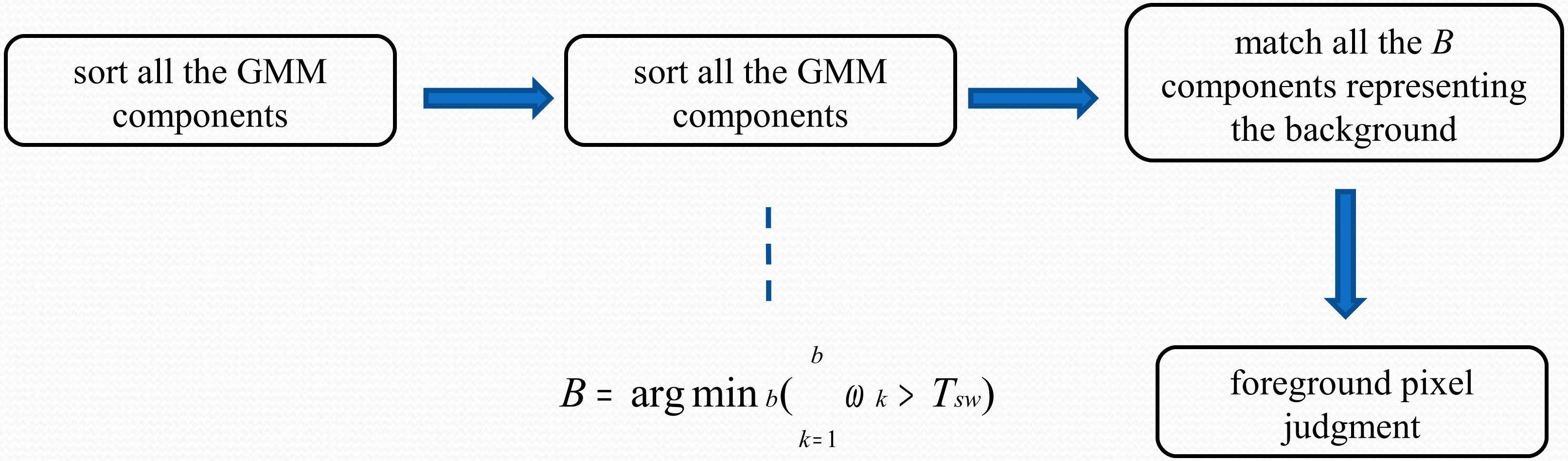
$$\rho = \alpha\eta(x_t \mid \mu_{i,t}, \sigma_{i,t}) \Rightarrow \rho_{i,t} = \frac{\alpha}{\omega_{i,t}}$$

The probability factor causes high computational cost

Improve the classical GMM

The idea – optimization by simplification

- Adapt a light-weight discriminant
  - The classical flow



$$B = \arg\min_b\left(\sum_{k=1}^{b} \omega_k > T_{sw}\right)$$

Improve the classical GMM

The idea – optimization by simplification

- Adapt a light-weight discriminant
  - The novel version for MST

calculate upon all the components with simplified → direct shadow judgment

an empirical constant

$$M_{i,t} = \begin{cases} 1, & |X_t - \mu_{i,t}| \le T_\sigma \\ 0, & otherwise \end{cases}$$

$D * \sigma_{i,t}$

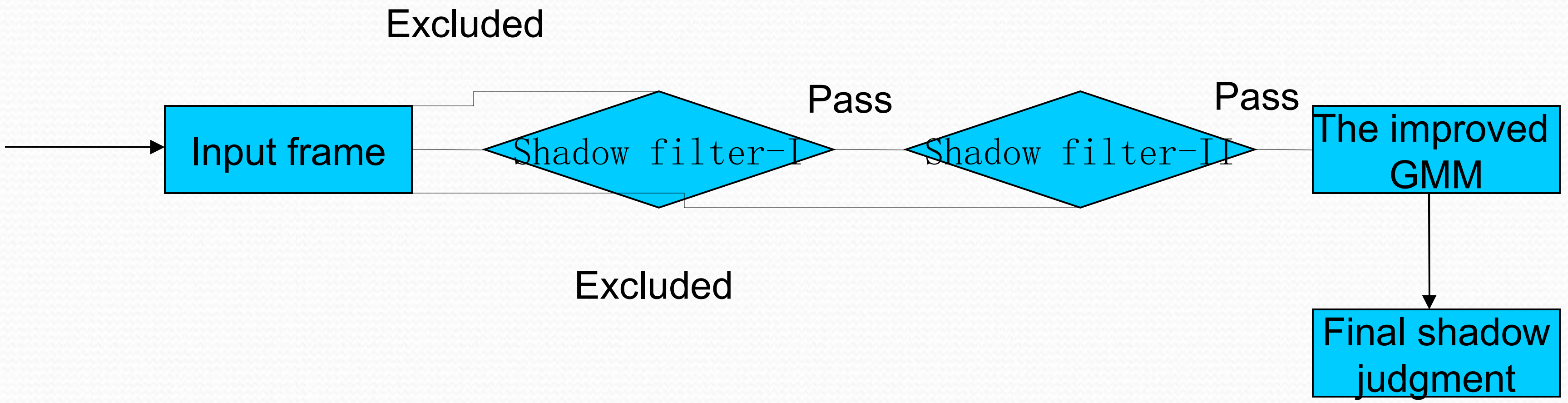
counterpart in the classical framework

$$SJ(x_t) = \begin{cases} 1, & \sum_{i \in \{k \mid M_{k,t}=1\}} \omega_{i,t} > T_{sw} \\ 0, & otherwise \end{cases}$$

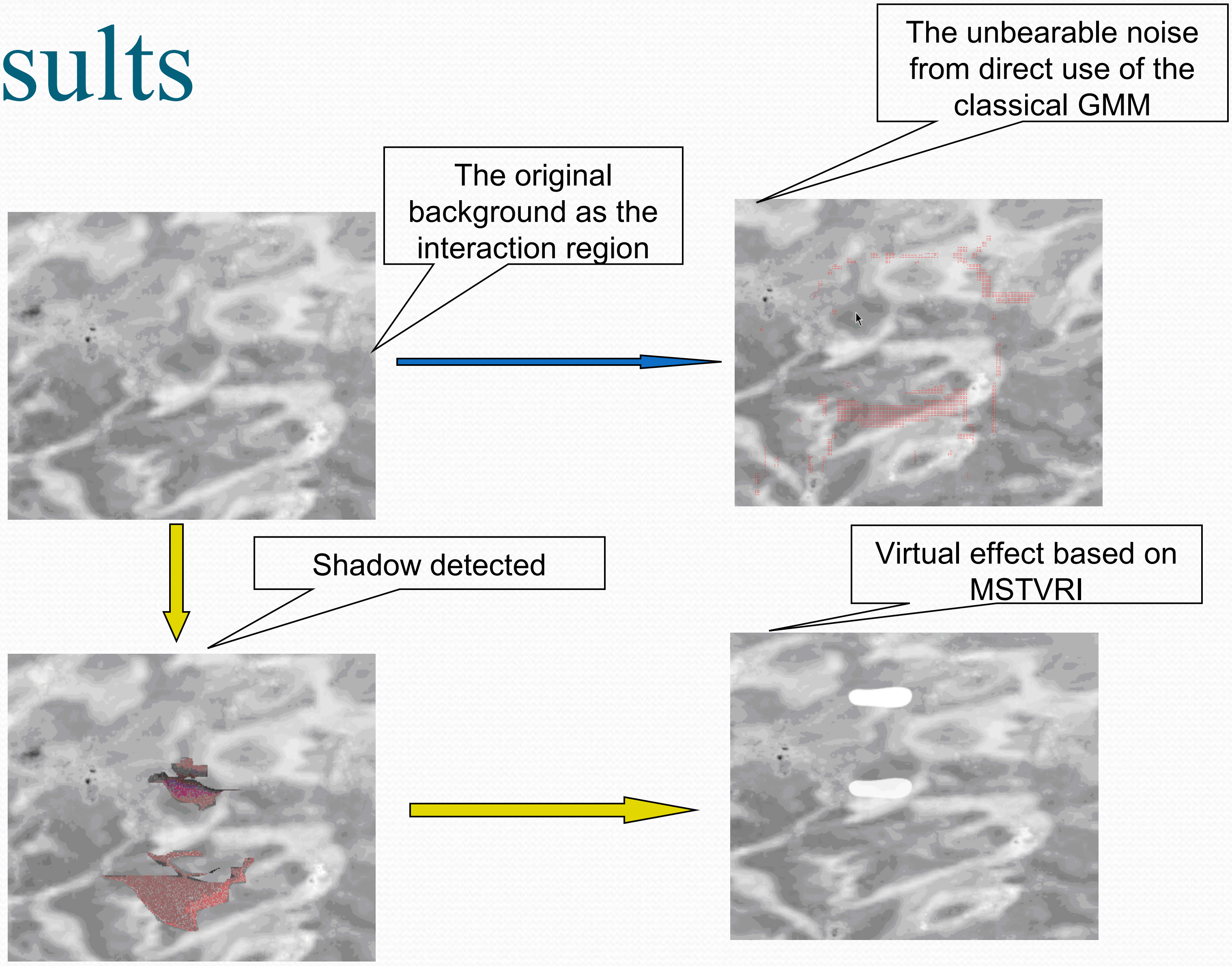
MSTVRI - overview

- The idea of MSTVRI (MST for VR Interaction)
  - Precede the final shadow judgment based on its motion feature with the two-step shadow discriminant

- The integral flow



Results


The original background as the interaction region
The unbearable noise from direct use of the classical GMM
Shadow detected
Virtual effect based on MSTVRI

Summary

- Work done
  - Introduce a shadow filter to preprocess the video frame to be detected so as to exclude pixels that is not probably in shadow area, thus save the otherwise subsequent extra process
  - Improve the classical GMM approach to motion detection by simplifying every possible items that is computationally expensive and thus cause high real-time performance loss

Summary

- Work to be done
  - The MSTVRI itself as an algorithm of moving shadow detection is fairly application-specific, far from being a optimal solution to general shadow detection
  - The VR interaction control would be limited while there are too many objects interacting with the video scene simultaneously, as cast interlaced moving shadows.


Questions?
...And
Answers!

Thanks！