Video Layer Extraction and Reconstruction

Sylvain Pelletier[1], Françoise Dibos[2] and Georges Koepfler[1]

[1]University Paris Descartes, MAP5 (UMR CNRS 8145)
[2]University Paris Nord, LAGA (UMR CNRS 7539)

MVA 2010-11

Plan:

Introduction

Layer decomposition

An image of a natural scene is obtained by projection of the 3D scene.
This is modeled by the superposition of several layers.
An occlusion appears if a layer is projected upon another.

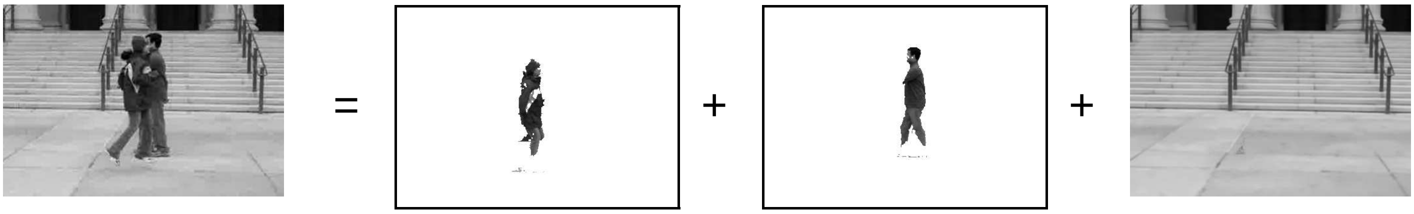
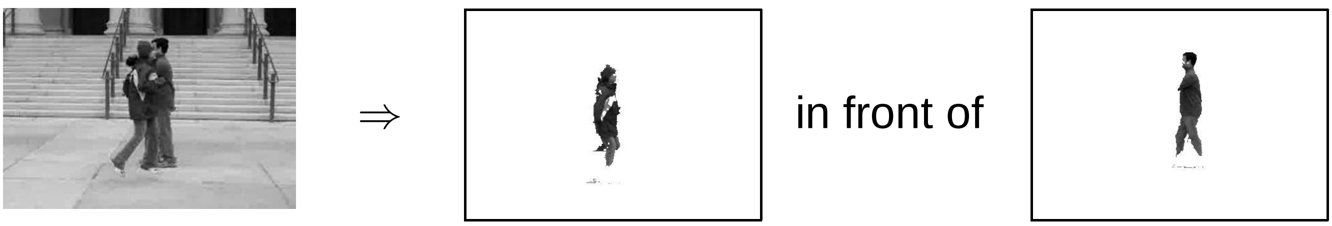


image = occluding person + occluded person + background

Introduction

Problem:

- extract and reconstruct layers from a video sequence;
- even if total occlusion occurs during several frames.



Example of **sequence** (occluding, occluded, hidden).

Layer model definition

Definition:

- image I defined on the domain $\mathcal{D}$;
- **layer** of 3d moving object: (Ω, o)
  - $\Omega \subset \mathcal{D}$ region where the object is projected if no occlusions;
  - o gray level function defined on Ω, giving the object's gray level.

Consider $N+1$-frame video sequence $(I_i)_{i=0,\dots,N}$ with:
a fixed background,
an occluded moving object and an occluding moving object.

Three layers:

- background layer $(\mathcal{D}, B)$;
- occluded object layer (Ω_i, o_{Ω_i});
- occluding object layer $(\mathcal{O}_i, o_{\mathcal{O}_i})$.

Layer model definition

According to the **layer model**, the image I_i at pixel $\mathbf{x} \in \mathcal{D}$ is given by:

$$I_i(\mathbf{x}) \approx \begin{cases} o_{\mathcal{O}_i}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{O}_i; \\ o_{\Omega_i}(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_i \setminus \mathcal{O}_i; \\ B(\mathbf{x}) & \text{else.} \end{cases}$$

The notation $\approx$ allows for unknown noise effects.

$$\begin{aligned} I_i(\mathbf{x}) \approx\ & o_{\mathcal{O}_i}(\mathbf{x})\mathbb{I}_{\mathcal{O}_i}(\mathbf{x}) \\ & + o_{\Omega_i}(\mathbf{x})\mathbb{I}_{\Omega_i}(\mathbf{x})\left(1 - \mathbb{I}_{\mathcal{O}_i}(\mathbf{x})\right) \\ & + B(\mathbf{x})\left(1 - \mathbb{I}_{\Omega_i}(\mathbf{x})\right)\left(1 - \mathbb{I}_{\mathcal{O}_i}(\mathbf{x})\right). \end{aligned}$$

Layer deformation model

Idea: 3d object motion model yields deformation functions,
T_i for the occluded object and T'_i for the occluding object.

T_i is such that:

pixel $\mathbf{x}_0 \in I_0$ and pixel $\mathbf{x}_i = T_i(\mathbf{x}_0) \in I_i$
are the projection of the same moving 3d point $\mathbf{X}$.

T_i and T'_i take care of the perspective deformation of the objects.

Layer deformation model

Assuming no occlusion in I_0: $I_i(\mathbf{x}) \approx \begin{cases} I_0(T_i'^{-1}(\mathbf{x})) & \text{if } T_i'^{-1}(\mathbf{x}) \in \mathcal{O}_0; \\ I_0(T_i^{-1}(\mathbf{x})) & \text{if } T_i^{-1}(\mathbf{x}) \in \Omega_0; \\ B(\mathbf{x}) & \text{else.} \end{cases}$

Image I_i with occlusion.Use T_i and Ω_0 (= the car).Occlusion of Ω_i can be restored.

Layer deformation model

Hypotheses:

- fixed camera and known background B;
- no occlusion in first image;
- moving objects are rigid;
- movement is a uniform translation in 3d space.

Thus:

- consider only short sequences;
- parametric deformation deduced from 3d motion and not from 2d.

Layer deformation model

Construction of **layer deformation function** T_i:

- moving 3d point $\mathbf{X}(t) = (X(t), Y(t), Z(t))$;
- uniform translation $\dot{\mathbf{X}}(t) = (A, B, C)$;
- projection $\mathbf{x}$ onto I_i by pin-hole camera model:

$$\mathbf{x}_i = \begin{bmatrix} x(t_i) \\ y(t_i) \end{bmatrix} = \begin{bmatrix} \frac{X_i}{Z_i} \\ \frac{Y_i}{Z_i} \end{bmatrix} = \frac{1}{ci\Delta t + 1} \left(\begin{bmatrix} x_0 \\ y_0 \end{bmatrix} + i\Delta t \begin{bmatrix} a \\ b \end{bmatrix} \right).$$

with $a = A/Z(0)$, $b = B/Z(0)$ and $c = C/Z(0)$.

$$\mathbf{x}_i = \frac{1}{1 + ci} (\mathbf{x}_0 + i\Delta t\mathbf{a}) = T_i(\mathbf{x}_0).$$

Variational model

$$I_i(\mathbf{x}) \approx \begin{cases} o_{\mathcal{O}_i}(\mathbf{x}) & \text{if } \mathbf{x} \in \mathcal{O}_i; \\ o_{\Omega_i}(\mathbf{x}) & \text{if } \mathbf{x} \in \Omega_i \setminus \mathcal{O}_i; \\ B(\mathbf{x}) & \text{else.} \end{cases} \quad \text{with } \mathcal{O}_i = T_i'(\mathcal{O}_0),\ \Omega_i = T_i(\Omega_0)$$

$$\text{and } \Omega_i \cap \mathcal{O}_i = T_i\left(\left\{\mathbf{x}_0 \in \Omega_0 \,/\, {T'}_i^{-1} T_i(\mathbf{x}_0) \in \mathcal{O}_0\right\}\right).$$

Terms used in the global energy:

- the static difference: $\Delta_i^0(\mathbf{x}) = I_i(\mathbf{x}) - B(\mathbf{x})$;
- the warp motion difference for the occluded object:

$$\Delta_i^1(\mathbf{x}) = I_i(\mathbf{x}) - I_0(T_i^{-1}\mathbf{x}) ;$$

- the warp motion difference for the occluding object:

$$\Delta_i^2(\mathbf{x}) = I_i(\mathbf{x}) - I_0(T_i'^{-1}\mathbf{x}) ;$$

- the boundary detection for the moving objects:

$$g\left(|\nabla I_0|\right) = 1/(1 + |\nabla I_0|^2) .$$

Variational model

Compute Ω_0, $\mathcal{O}_0$ and (a, b, c, a', b', c') from the sequence.

Energy

$$\begin{aligned} E &= \sum_{i>0} \int_{\mathcal{O}_i} \rho\left(\Delta_i^2(\mathbf{x})\right) d\mathbf{x} \\ &+ \sum_{i>0} \int_{\Omega_i \setminus \mathcal{O}_i} \rho\left(\Delta_i^1(\mathbf{x})\right) d\mathbf{x} \\ &+ \sum_{i>0} \int_{D \setminus (\mathcal{O}_i \cup \Omega_i)} \rho\left(\Delta_i^0(\mathbf{x})\right) d\mathbf{x} \\ &+ \lambda \int_{\partial\Omega_0} g\left(|\nabla I_0|\right) ds + \lambda' \int_{\partial\mathcal{O}_0} g\left(|\nabla I_0|\right) ds , \end{aligned}$$

with robust estimator, *e.g.* $\rho(s) = \sqrt{\epsilon^2 + s^2}$.

Energy (after change of variables)

Variational model: Minimization

- Initialisation
  - get regions $\mathcal{R}_1$ and $\mathcal{R}_2$: compare I_0 and B;
  - parameters (a, b, c, a', b', c'):
    use parametric optical flow between I_0 and I_1.
- Depth order:

  test $(\Omega_0, \mathcal{O}_0) = (\mathcal{R}_1, \mathcal{R}_2)$ versus $(\Omega_0, \mathcal{O}_0) = (\mathcal{R}_2, \mathcal{R}_1)$
  and keep the one with the lowest energy E.
- Iterate:
  - Minimize on regions Ω_0 and $\mathcal{O}_0$ with I.C.M. ;
  - Minimize on parameters (a, b, c, a', b', c') with simplex method.

Variational model: Minimization on regions

$$E(\Omega_0, \mathcal{O}_0) = \int_{\mathcal{O}_0} \sum_{i>0} f_i'(\mathbf{x}_0) d\mathbf{x}_0 + \int_{\Omega_0} \sum_{i>0} f_i(\mathbf{x}_0) \mathcal{V}_i(\mathbf{x}_0, \mathcal{O}_i) + \lambda \int_{\partial\Omega_0} g(|\nabla I_0|)\, ds + \lambda' \int_{\partial\mathcal{O}_0} g(|\nabla I_0|)\, ds,$$

with $\mathcal{V}_i(\mathbf{x}_0, \mathcal{O}_i) = 1$, if pixel $\mathbf{x}_0$ is visible in I_i.

ICM

For two regions: compute $\phi : \hat{D} \to \{-1, 0, 1\}$,
such that $\Omega_0 = \{\phi = 1\}$ and $\mathcal{O}_0 = \{\phi = -1\}$.

Variational model: Minimization on regions

$$E(\Omega_0,\mathcal{O}_0) = \int_{\mathcal{O}_0} \sum_{i>0} f'(\mathbf{x}_0) d\mathbf{x}_0 + \int_{\Omega_0} \sum_{i>0} f_i(\mathbf{x}_0)\mathcal{V}_i(\mathbf{x}_0,\mathcal{O}_i)$$
$$+ \lambda \int_{\partial\Omega_0} g\left(|\nabla I_0|\right) ds + \lambda' \int_{\partial\mathcal{O}_0} g\left(|\nabla I_0|\right) ds \, ,$$

$$\hat{E}(\phi) = \sum_{s\in\hat{D}} \hat{F}'(s,\phi_s) + \sum_{s\in\hat{D}} \sum_{i} \hat{F}_i\left(s,\phi_s,\phi_{O_i}\right) + \sum_{s\in\hat{D}} \hat{L}\left(s,\phi_s,\{\phi_t\}_{s\sim t}\right) \, ,$$

with O_i the set of pixels occluding pixel s in I_i.

$\hat{E}(\phi_s)$ depends on:

- ϕ_s, other pixels are fixed;
- $\{\phi_t\}_{s\sim t}$, the neighborhood of s (length term);
- $(O_i)_{i=1,\ldots,N}$, the occluding pixels (non local term).

Variational model: Minimization on parameters

$$E(a,b,c,a',b',c') = \int_{\mathcal{O}_0} \sum_{i>0} f_i'(a',b',c',\mathbf{x}_0) d\mathbf{x}_0$$
$$+ \int_{\Omega_0} \sum_{i>0} f_i(a,b,c,\mathbf{x}_0)\mathcal{V}_i(\mathbf{x}_0,\mathcal{O}_i) d\mathbf{x}_0 \, ,$$

with $\mathcal{V}_i(\mathbf{x}_0,\mathcal{O}_i) = 1$, if pixel $\mathbf{x}_0$ is visible in I_i.

Simplex Method

Apply the method to the continuous parameters (a,b,c,a',b',c').

Results on synthetic sequence

original sequence

restored sequence

Results on real sequences

Office Sequence

Outdoor Sequence.

Conclusion

- extract layers of moving objects from a sequence;
- reconstruct occluded parts;
- 3d motion model takes into account deformations;
- variational formulation allows to recover depth information.
- restore sequences, inversion of layers.