



# Object Tracking using Joint Visible and Thermal Infrared Video Sequences

Internal report

Renaud Péteri and Ondřej Šiler

Laboratoire de **M**athématiques **I**mage et **A**pplications  
Université de La Rochelle  
Avenue Michel Crépeau, 17042 La Rochelle Cedex  
FRANCE

2009

### Abstract

Particle filter methods based on color distribution can be used to track non-rigid moving objects in color videos. They are robust in case of noise or partial occlusions. However, using particle filters on color videos is sensitive to changes in the lighting conditions of the scene. The use of thermal infrared image sequences can help the tracking process, as thermal infrared imagery is not sensitive to lighting conditions. This paper presents a particle filter based method for object tracking using automatic cooperation between the color and the infrared modalities. As the infrared and the visible image sequences are acquired with different cameras, a pre-step is spatio-temporal registration. After spatio-temporal registration, the proposed method is able to continuously track the target despite difficult conditions appearing in one of the modality. Our cooperative tracking method is successfully applied on several experimental datasets. Different test sequences are presented, including tracking in the visible video with the help of the infrared modality, or tracking in the infrared with the help of the visible modality. Comments and future prospects raised by this method are finally described.

**keywords:** tracking, particle filtering, thermal infrared, color distribution, image sequences

# Contents

Introduction . . . . .	2
Particle Filtering . . . . .	3
Introduction . . . . .	3
The CONDENSATION algorithm . . . . .	3
Using visual and thermal infrared modalities for tracking . . . . .	4
Spatio-temporal registration . . . . .	4
Temporal registration . . . . .	4
Spatial registration . . . . .	4
Applying particle filtering on color and thermal imagery . . . . .	6
State space . . . . .	6
Dynamical model . . . . .	6
Observation models . . . . .	7
Weighting of the samples . . . . .	7
Proposed method for cooperative tracking using visual and IR information . . . . .	8
Experiments . . . . .	9
Tracking in the IR modality when target is lost in the visible one Introduction . . . . .	9
Description of the different steps . . . . .	9
Spatio-temporal registration . . . . .	9
Tracking in the visible modality . . . . .	9
Tracking using both modalities . . . . .	10
Comments . . . . .	10
Tracking in the visible modality when target is lost in the IR one	13
Conclusion and prospects . . . . .	16

## Introduction

Object tracking is a very intense research area in computer vision due to the wide range of concerned domains. In surveillance applications, tracking people is required to distinguish normal behaviors from suspicious ones[9]. In traffic surveillance, it allows to evaluate traffic density and car speed in order to detect accidents[8]. Tracking is used in motion analysis to help reeducation of disabled people or for the training of professional athletes. In mobile robot applications, tracking is needed to handle occlusions and disappeared objects [5].

Many techniques have already been proposed, and can be divided in two groups[7]. On the one hand, *bottom-up* approaches consist in partitioning the image in consistent regions. These regions are then matched based on a similarity measure. On the other hand, *top-down* techniques generate hypotheses representing hypothetical target states and check them using measurement extracted from images. Mean Shift, Bayesian filters like Kalman and particle filters belong to these techniques. Kalman filter propagates and updates the mean and the covariance of the posterior distribution in linear and Gaussian estimation problems. However, in real world applications characterized by strong clutter, the posterior is multimodal and can be approached by particle filter. Particle filter methods based on color distribution have been proven to be to very efficient for tracking non-rigid moving objects in color videos (see *Nummiaro et al.*[7]; *Bichot et al.*[2]). They are indeed robust in case of noise or partial occlusions. However, most of tracking methods are sensitive to changes in the lighting conditions of the scene (for instance a person crossing a shadowed area). When available, thermal infrared cameras can help the tracking process in case of difficult lighting conditions. Thermal infrared imagery is indeed nearly invariant to illumination changes.

The main contribution of this paper is to jointly use color and thermal infrared modalities during tracking. Our cooperation method is based on particle filtering, running in a cooperative way on both visible and infrared sequences: if a difficulty arises in one of the modality, our algorithm tries to track in the second one, till it is again able to track in the first modality.

In section of this paper, particle filtering is presented in the general case. As the infrared and the visible image sequences are acquired with different cameras, a spatio-temporal registration is performed as a pre-step. The application of particle filtering for color and thermal infrared images is exposed, and a method for cooperative tracking between this two modalities is presented (section ). The algorithm is then successfully applied on several experimental datasets. Different test sequences are presented, including tracking in the visible video with the help of the infrared modality, or tracking in the infrared with the help of the visible modality (section ). Comments and future prospects raised by this method are finally described (section ).

# Particle Filtering

## Introduction

Particle filtering is very successful for non-linear and non-Gaussian estimation of target state at time  $t$ , denoted by  $X_t$ , conditionally to measurements  $Z_{1:t}$  from first time to time  $t$ . This technique is used to implement a recursive Bayesian filter and is reliable in cases of clutter and during occlusions. The key idea is to represent a required posterior density function by a set of random samples  $\{s_t^{(n)}, n = 1 \dots N\}$  with associated weights  $\{\pi_t^{(n)}, n = 1 \dots N\}$  and to compute estimates based on these samples and weights thanks to

$$p(X_t|Z_{1:t}) = \sum_{n=1}^N \pi_t^{(n)} \delta(x - s_t^{(n)}) \quad (1)$$

$s_t^{(n)}$  is a hypothetical state of target and  $\pi_t^{(n)}$  is proportional to the probability of the state  $s_t^{(n)}$ . The set of samples is propagated through time in two steps. Firstly, samples are selected proportionally to their weights: samples with high weights are duplicated, whereas those with low weights are eliminated (resampling step). Sample are then propagated and weighted according to an importance density that describes regions of interest in the state space.

As the number of samples becomes very large, this characterization becomes an equivalent representation of the usual functional description of the posterior probability density function  $p(X_t|Z_{1:t})$  (i. e. pdf) and the particle filter approaches the optimal Bayesian estimate.

There are many particle filtering methods which differ on the resampling technique and the choice of the importance function. Reviews are proposed in Arulampalam *et al.* work [1] and in Hue thesis[3]. In the next section, we focus on the specific filter called CONDENSATION [4] in which the importance density is the dynamic model.

## The Condensation algorithm

In the CONDENSATION [4] method, a particle  $s_{t-1}^{(n)}$  is drawn randomly from  $p(X_{t-1}|Z_{1:t-1})$  by choosing it, with probability  $\pi_{t-1}^{(n)}$ , from the set of  $N$  samples at time  $t-1$ . Next draw  $s_t^{(n)}$  randomly from  $p(X_t|X_{t-1} = s_{t-1}^{(n)})$ , one time-step of the dynamical model, starting from  $x_{t-1} = s_{t-1}^{(n)}$ . A value  $s_t^{(n)}$  chosen in this way is a fair sample from  $p(X_t|Z_{1:t-1})$ . It can then be retained as a pair  $(s_t^{(n)}, \pi_t^{(n)})$  for the  $N$ -set at time  $t$ , where  $\pi_t^{(n)} = p(Z_t|X_t = s_t^{(n)})$ .

The flow chart and the algorithm of this method are respectively given in figure 1 and algorithm 1.

In practice, random variables can be generated efficiently, using binary search, if we store cumulative probabilities  $c_{t-1}^{(n)}$  as shown in algorithm 1 instead of storing probabilities  $\pi_{t-1}^{(n)}$ .

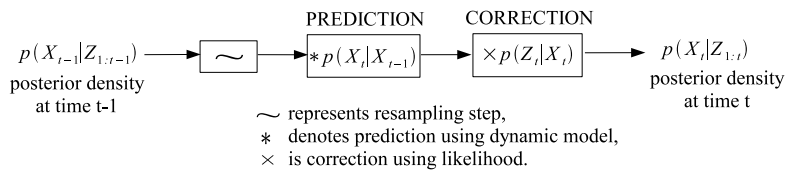


Figure 1: A time step in the CONDENSATION algorithm as described in MacCormick and Isard work [6].

## Using visual and thermal infrared modalities for tracking

### Spatio-temporal registration

Video registration is a necessary pre-step for being able to integrate data obtained from different cameras. Spatial registration is the process of transforming the different datasets into one coordinate system. Temporal registration contemporizes corresponding events. In our case, a thermal infrared and a color camera are recording the scene where the target moves. Both cameras are located during experiments as closed to each other as possible, which enables to approximate the distortion between the two images by an affine transformation. We have chosen to first register temporally the image sequences, and then spatially.

#### Temporal registration

The two videos have to be temporally synchronized to permit to switch at precise moments between the visible and IR (infrared) modalities. Video sequences are arranged in order to have the same frequency (frame per second - fps). One keeps every  $n^{\text{th}}$  frame from the video with the highest fps, where  $n$  is the ratio between frequencies. In order to find the beginning and the end of each sequence, the essential task is to identify a frame with corresponding events in both sequences. Another possibility for synchronization is to use a digital timer visible in both sequences.

#### Spatial registration

Because of the acquisition conditions, affine transformation is a reasonable assumption for the registration precision we need for tracking. An affine transformation is a superposition of elementar linear transformations (translation, rotation and scaling), and can be expressed by only one transformation matrix.

The affine transformation from a point  $\mathbf{X}$  to a point  $\mathbf{X}'$  is given by:

$$\mathbf{X}' = \mathbf{M} \mathbf{X} \quad (2)$$

---

**Algorithm 1:** The CONDENSATION algorithm
 

---

**Input:**  $\{(s_{t-1}^{(n)}, \pi_{t-1}^{(n)}), n = 1 \dots N\}$  the particle set  $s_{t-1}^{(n)}$  at previous time step and their associated weights  $\pi_{t-1}^{(n)}$

$N$  : total number of particles

**Output:**  $\{(s_t^{(n)}, \pi_t^{(n)}), n = 1 \dots N\}$  the particle set at current time step  $t$

- Resampling
  - Generate a random number  $r \in [0, 1]$ , uniformly distributed.
  - Find, by binary subdivision on  $m$ , the smallest  $m$  for which  $c_{t-1}^{(m)} \geq r$ .

- Prediction
  - Draw a random variable  $s_t^{(n)}$  from the dynamical density  $p(X_t | X_{t-1} = s_{t-1}^{(n)})$ .

- Correction
  - Update weights  $\pi_t^{(n)} \propto p(Z_t | X_t = s_t^{(n)})$  and normalise so that  $\sum_{n=1}^N \pi_t^{(n)} = 1$
  - Store together samples with cumulative probabilities as  $(s_t^{(n)}, \pi_t^{(n)}, c_t^{(n)})$  where

$$\begin{aligned} c_t^{(0)} &= 0, \\ c_t^{(n)} &= c_t^{(n-1)} + \pi_t^{(n)} \quad (n = 1, \dots, N) \end{aligned}$$

- Estimate target state using the mean :  $E[X_t] = \sum_{n=1}^N \pi_t^{(n)} s_t^{(n)}$
- 

where the affine transformation matrix  $M$  is

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ 0 & 0 & 1 \end{pmatrix}.$$

and the point coordinates are considered as vectors extended by the normalizing constant 1 (a point  $X$  with coordinates  $X_x$  and  $X_y$  is represented by the triplet  $\mathbf{X} = (X_x, X_y, 1)^T$ ).

There are six unknown parameters in matrix  $M$ , so three control points in both images are needed to determine the image transformation. Using this control points, the coefficients  $m_{ij}$  are obtained by inverting matrix  $M$  and solving this linear system.

For applying this affine transformation to the image to be registered, a new image with required size is created. Then for each pixel in the new image we

compute the corresponding coordinates in the original image using the inverse transform  $M^{-1}$ . In most cases, these coordinates are not integers, so values of pixels in the new image are obtained by linear interpolation of neighboring pixels in the original image.

## Applying particle filtering on color and thermal imagery

In order to apply the particle filter on our visible and infrared image sequences, we have to define the state space, the dynamical and the observation models.

### State space

State space represents a set of all unknown information characterizing a tracked object in time  $t$ . In our case, we have defined the state space as an upright rectangular region with half width and length  $H_x$  and  $H_y$ . State vector is then:

$$X_t = \{x_t, y_t, H_{x_t}, H_{y_t}\}, \quad (3)$$

where  $x_t$  and  $y_t$  are the coordinates of the center of the rectangle,  $H_{x_t}$  and  $H_{y_t}$  are the half width and length of the rectangle.

### Dynamical model

The dynamical model describes the evolution of the process. It relies on positions and velocities estimated in the previous frames. Velocity is updated according to the difference of mean target position in the last frame and the frame before last as follows:

$$\dot{x}_t = \bar{x}_{t-1} - \bar{x}_{t-2}, \quad (4)$$

$$\dot{y}_t = \bar{y}_{t-1} - \bar{y}_{t-2}, \quad (5)$$

where  $\bar{x}$  and  $\bar{y}$  is the mean value of the  $x$  and  $y$  coordinate respectively. The new length and width of the rectangle are either the same as previous, or  $\pm 10\%$  chosen randomly with normal distribution. Finally, the model is given by

$$x_t = x_{t-1} + \dot{x}_{t-1} + \sigma\omega_t, \quad (6)$$

$$y_t = y_{t-1} + \dot{y}_{t-1} + \sigma\omega_t, \quad (7)$$

$$H_{x_t} = \begin{cases} H_{x_{t-1}} \\ H_{x_{t-1}} \pm 10\%H_{x_{t-1}} \end{cases}, \quad (8)$$

$$H_{y_t} = \begin{cases} H_{y_{t-1}} \\ H_{y_{t-1}} \pm 10\%H_{y_{t-1}} \end{cases}, \quad (9)$$

where  $\sigma$  is the standard deviation of dynamical noise and  $\omega$  is unit zero-mean Gaussian noise.

## Observation models

We want to apply the particle filter on two modalities (visible and thermal infrared), thus we need two different observation models. Color distributions are used as target models in the color video sequences and thermal distribution in infrared sequences. Color and temperature histograms are chosen because they achieve robustness against non-rigidity, rotation and partial occlusions.

### Observation model for color image

Real world color video sequences are characterized by strong lighting changes. Hence, in many applications, tracking uses rgb normalized space to reduce the effect of these changes. We have chosen the observation model similar to the work of Nummiaro *et al.*[7]. The distributions are determined inside an upright rectangular region with half length and width  $H_x$  and  $H_y$ . To increase the reliability of the distribution when boundary pixels belong to the background or get occluded, smaller weights are assigned to the pixels that are further away from the region center.

### Observation model for thermal infrared image

An infrared camera measures and images the emitted infrared radiation from an object. If properly calibrated, a thermal infrared camera can give a measurement of the temperature of the object, as radiation is a function of object surface temperature.

On infrared images, our observation model is computed using temperature histogram instead of color histogram.

## Weighting of the samples

To weight the samples at the correction step of the CONDENSATION method, we compare target color distribution model  $q$  with the distributions  $p_{s_t^{(n)}}$  evaluated on the rectangular regions defined by hypothesis  $s_t^{(n)}$  (called candidate distribution). The comparison is performed using the Bhattacharyya distance  $D(\cdot)$ . The weight is expressed by

$$\pi_t^{(n)} \propto \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{D^2(p_{s_t^{(n)}}, q)}{2\sigma^2}\right) \quad (10)$$

where  $\sigma$  represents variance of measurement noise.

During filtering, samples with a high weight may be chosen several times, leading to identical copies, which are diffused by dynamical noise, while others with relatively low weights may not be chosen at all.

Finally, the mean state of the object is estimated at time  $t$  by

$$E[X_t] = \sum_{n=1}^N \pi_t^{(n)} s_t^{(n)} \quad (11)$$

## Proposed method for cooperative tracking using visual and IR information

We will assume that during the whole sequence, the object is trackable in at least one modality. The goal of the proposed method is to continuously track the target despite difficult conditions appearing temporally in one of the modality (like change of lighting, occlusion in one modality. . .). Suppose we want to track an object in a color video: if the target is being lost, we then switch to the infrared sequence, where tracking is supposed possible. We use the Bhattacharyya distance of color distributions as a measure of similarity between candidates and a reference model. When the target disappears, the color distribution changes and the Bhattacharyya distance between the reference model and the estimated position histograms increases. If the distance exceeds a given threshold, the target is considered to be lost. We then begin to track the object in the IR video. Starting position for tracking is set to previous correctly estimated position in color video that is, thanks to video registration, at the same coordinate system as IR video. Tracking in IR video is then performed while tracking in the visible is periodically tested. If the Bhattacharyya distance is less than the threshold, the algorithm stops tracking in IR and continues only in visible video. The diagram of this method is shown on figure 2. The same principle also stands when infrared video is the first modality and color video the second.

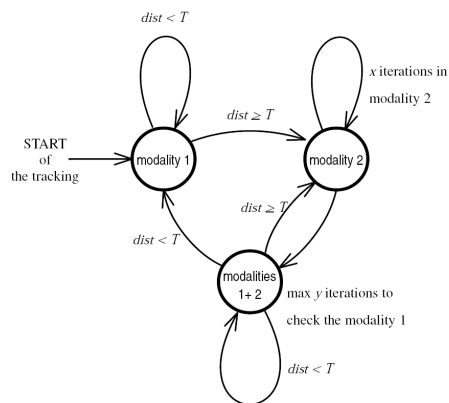


Figure 2: State diagram of the cooperative tracking method: tracking starts in the *modality 1*. After every processed frame, the Bhattacharyya distance  $dist$  is evaluated. If it is greater than the threshold  $T$  relative to modality 1, then tracking continues for  $x$  frames in *modality 2*. Both videos are processed (*modalities 1 + 2*) for a maximum of  $y$  frames. During this trial period, the Bhattacharyya distance is again computed in modality 1. If above the threshold  $T$ , then the tracking comes back to *modality 1*, otherwise it returns to *modality 2*.

## Experiments

### Tracking in the IR modality when target is lost in the visible one

#### Introduction

The following experiment illustrates how thermal infrared sequences can be used when tracking is temporally lost in the visible modality. Our thermal infrared camera is a ThermoVision A40M - Researcher. This camera delivers high-resolution images ( $320 \times 240$  pixels) at a refresh rate of 50/60 Hz. It gives measurement in the spectral range from 7.5 to  $13 \mu\text{m}$ . This wavelength range includes the thermal radiation of objects around the human body temperature ( $\approx 310 \text{ K}$ ).

The aim of the described experiment was to track a coffee cup filled with hot water. Temperature of water was around 330 K, thus distinguishable in the IR. The cameras were located at about 4 meter far from the assumed target trajectory, with a mutual distance of about 20 centimeters. IR video was recorded at a frequency of 60 fps and with a resolution of  $320 \times 240$  pixels. Visible video was recorded at a frequency of 30 fps and with a resolution of  $640 \times 480$  pixels. The experiment was the following: at the beginning, lamps were switched on and created almost constant lighting. Under these conditions, tracking in the visible video with particle filtering was possible. After some time, lamps were switched off. The partially obscured room was then too dark to keep on tracking in the visible video. After a few seconds, the lamps were switched on again.

#### Description of the different steps

##### Spatio-temporal registration

For temporal synchronization, every second frame was discarded from the IR video to align fps rates of both videos. A frame with a rapid movement of a chair hit by the the person handling the cup was picked as a reference event in both sequences.

For spatial registration, two of the control points were chosen on the target trajectory and one on an object in the motion plane. The spatial image registration of the visible video on the IR video is shown on Figure 3.

##### Tracking in the visible modality

When the target is tracked only in the visible video, it is lost after switching the lights off (Fig. 4). Because of the weak illumination, the target is indeed not distinguishable from its environment and too far from its original model.

The change in the lighting conditions can be observed on the Bhattacharyya distance graph (Fig. 5): it generates a rapid growth of the Bhattacharyya distance. After relighting, the distance falls again below the empirically defined



Figure 3: Spatial registration of the visible video (left) on the IR video (right). The red rectangle is the field of view common to both sequences.

threshold. Particle filter seeks neighboring areas and searches for color distribution as similar as possible to the target model. After switch off and on the lights, the estimation of target position diverges from the real target position and does not recover: it stabilizes on a poster in the background.

#### Tracking using both modalities

As described in the previous section, tracking only in the visible modality fails. Our cooperative tracking algorithm is then applied on both modalities (Fig. 6). The tracking starts in the visible video, as for the previous example. After reaching the Bhattacharyya distance threshold when the lights are off, it switches to IR video, where the target can be followed. The algorithm periodically tries to track in the visible for several frames, but as the room is still in the dark, the Bhattacharyya distance remains too high. One recalls that, for each try in the visible sequence, the tracking is launched from the last estimated position of the IR sequence. After relighting, the Bhattacharyya distance in color video falls below the defined threshold. In the next trial stage, the program switches back to visible video and continues tracking. The same process happens a second time when lights are again switched off and on (see figure 5).

#### Comments

This experiment has shown the relevance of the proposed cooperative method. Tracking in the visible modality can not be performed under difficult conditions, as colors are not sufficiently distinguishable after switching the lights off (even while normalized color space is used). Cooperative tracking with infrared video allows to follow the target during the whole record. The use of the infrared modality enables to overcome temporary occlusion (created by the lighting conditions). The tracking is performed in the visible modality once the reference model is again valid.

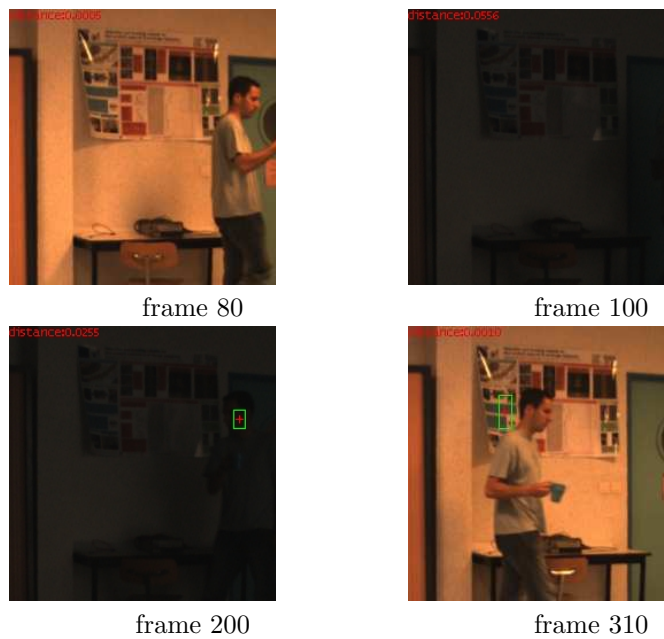


Figure 4: The “coffee cup” sequence: tracking in the visible video. The estimated position is displayed with a red cross and a green rectangle. Till frame 87 it is possible to track easily because lights are switched on. In frame 100 and 200, the estimated position has definitely left the coffee cup and follows the head of the person. In frame 310, lamps are relighted but the particle filter is not able to relocate the correct position, and stabilizes on a poster in the background.

The next experiment illustrates conversely how visible modality can overcome occlusions appearing when the tracking is performed in the IR modality.

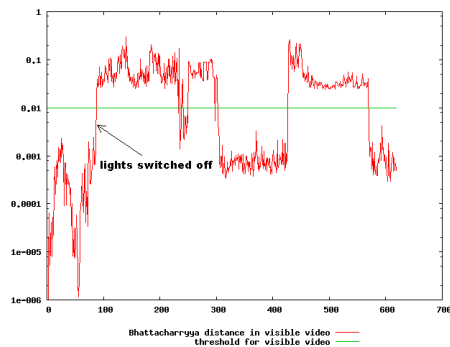


Figure 5: The “coffee cup” sequence: Bhattacharyya distance graph. In frame 87, the lights are switched off and the Bhattacharyya distance grows rapidly. In frame 240, lamps are relighted but the particle filter has already lost the target. Afterwards lights are again switched off in frame 428 and on in frame 574.



Figure 6: The “coffee cup” sequence: cooperative tracking using visible and IR video. Frames in grayscale (100 and 200) originate from the IR video while frames 80 and 310 originate from the visible camera. IR video assumes tracking after switching off the lights and continues until lights are switched on. Then the visible video assumes tracking again.

## Tracking in the visible modality when target is lost in the IR one

Some materials do not transmit electromagnetic waves in particular wavelength regions. Our thermal infrared camera gives measures in the wavelength range from 7.5 to 13  $\mu\text{m}$ . Glass is opaque in this part of the electromagnetic spectrum. Given that, it is not possible to see and track objects behind windows, glasses, windscreens etc.

Here, we are testing the opposite case compared to the previous experiment: tracking in the infrared video till the target is lost. As mentioned, target motion behind an IR obstacle - here, a window - makes tracking impossible in the IR modality.

In our case, we are tracking a bottle of frozen water at a temperature of approximately  $-5\text{ }^{\circ}\text{C}$ , which is moving partially behind a window. While behind the window, the bottle is not discernible in IR video and the target is lost (Fig. 7). The Bhattacharyya distance goes above the selected threshold (Fig. 8). The particle filter detects the bottle when it reappears. The reason is that the bottle is the only object in the scene with that particular temperature distribution and size. Comparable behavior arises several times in the sequence. Figure 9 shows tracking of the target using both modalities during the experiment. The cooperative method swaps between IR and visible modalities when the bottle is above/behind the window.

This experiment demonstrates that it is possible to use the cooperative method in the other way, when visible video helps the tracking in an IR video. This situation can occur for instance when we want to track with an IR camera a person getting in a car: the car screen is opaque to IR wavelength.

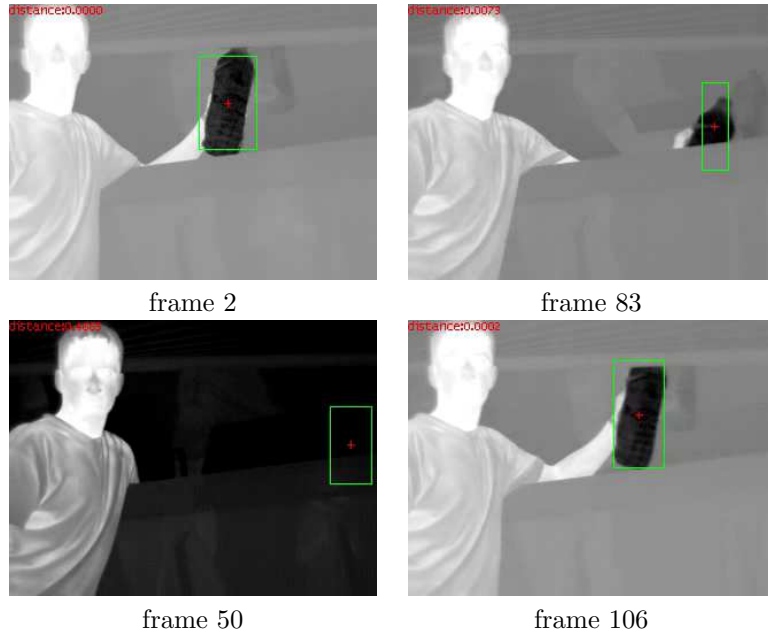


Figure 7: The “cold bottle” sequence: tracking in IR video. It is possible to track when the bottle is above the edge of the window edge (up to frame 32). When behind the window, it disappears from the IR video (frames 33 to 79). Then the particle filter recovers the target and tracks again the bottle. Same action repeats four times in the sequence.

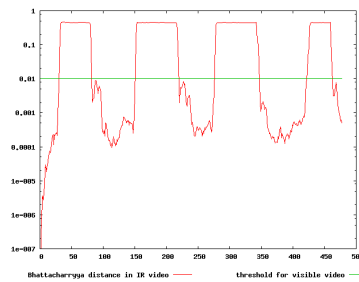


Figure 8: The “cold bottle” sequence: Bhattacharyya distance graph. From the plot, it is clear when the bottle is hidden and reappears – Bhattacharyya distance steeply grows (frames 32, 152, 277 and 424) and falls (frames 84, 221, 348 and 464) in that particular moments.

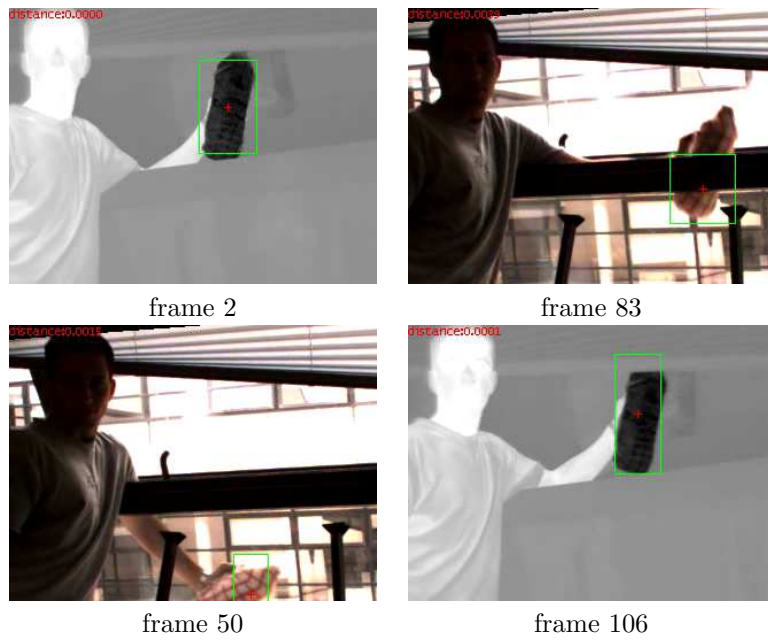


Figure 9: The “cold bottle” sequence: cooperative tracking in IR and visible modalities. This figure shows frames at same time as Fig. 7. When it was impossible to track in IR behind the window, color video helps to keep track of the target (frames 50 and 83).

## Conclusion and prospects

In this paper, a method for cooperative tracking using color and infrared image sequences has been proposed. A particle filter is used as a tracking algorithm in both modalities. It searches for color or temperature distributions similar to the target model. The relevance of this method has been shown on two examples, where one modality assumed the tracking during an occlusion occurring in the other modality.

Due to cost reasons, visible cameras have often a better spatial resolution than thermal infrared ones. An application of this work could be to track in the visible range for a precise spatial segmentation, and when the target is temporally lost, keeping track of the target in the infrared sequence, even if it implies in return a lost in the spatial precision.

Current works aim at updating the model of the object during the tracking. Here, we have supposed that the object model was the same before and after the occlusions. If the model changes, the difficulty is to set the threshold on the Bhattacharyya distance in order to be able to distinguish between a real change in the model and an occurring occlusion.

On infrared images, observation model based only on temperature histogram may be unable to distinguish two objects crossing each other and having the same temperature (for instance two cars on a road crossing each other). Adding motion information can help to overcome this case. An importance density function based on motion has already been proposed for color sequences in our previous works[2]. We are currently integrating motion information in the particle filters in order to increase the tracking robustness in the thermal infrared sequences.

# Bibliography

- [1] S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for on-line non-linear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50(2):174–188, 2002.
- [2] E. Bichot, L. Mascarilla, and P. Courtellemont. Fish tracking by combining motion based segmentation and particle filtering. In *Visual Communication and Image Processing*, volume 16077-28 of *Proceedings of SPIE-IS&T Electronic Imaging*, pages 1–10, 2006.
- [3] C. Hue. *Méthodes séquentielles de Monte Carlo pour le filtrage non linéaire multi-objets dans un environnement bruité. Applications au pistage multi-cibles et à la trajectographie d'entités dans des séquences d'images 2D*. PhD thesis, Université de Rennes, 2003.
- [4] M. Isard. *Visual Motion Analysis by Probabilistic Propagation of Conditional Density*. PhD thesis, Oxford University, 1998.
- [5] E. Koller-Meier and F. Ade. Tracking multiple objects using the condensation algorithm. *Journal of Robotics and Autonomous Systems*, 34:93–105, 2001.
- [6] J. MacCormick and M. Isard. Partitioned sampling, articulated objects and interface-quality hand tracking. In *Proceedings European Conference on Computer Vision*, volume 2, pages 3–19, 2000.
- [7] K. Nummiaro, E. Koller-Meier, and L. V. Gool. An adaptive color-based particle filter. *Image and Vision Computing*, 21(1):99–110, 2003.
- [8] J. Rittscher, J. Kato, S. Joga, and A. Blake. A probabilistic background model for tracking. In *Proc. European Conf. Computer Vision*, volume 2, pages 336–350, 2000.
- [9] N.T. Siebel. *Design and Implementation of People Tracking Algorithms for Visual Surveillance Applications*. PhD thesis, Department of Computer Science, The University of Reading, 2003.