An Autonomous Robotized System for a Thermographic Camera

Alberto Pretto1,3, Emanuele Menegatti1,3, Paolo Bison2, Ermanno Grinzato2, Gianluca Cadelano2 and Enrico Pagello1,3

1 University of Padova, Dep. of Information Engineering, Via Gradenigo 6/B - Padova (Italy)
2 CNR-ITC, Corso Stati Uniti 4 - Padova (Italy) http://www.itc.cnr.it
3 IT+Robotics Srl, Contrà Valmerlara 21 - Vicenza (Italy) http://www.it-robotics.it

Abstract

This paper presents a new approach to automatically monitor an indoor environment on thermodynamic basis. It uses temperature as the driving parameter and is especially suited for comfort analysis or evaluation of moisture. The system measures all fundamental environment parameters (e.g., air temperature, relative humidity and air speed) by imaging with a thermal camera (IR camera) a set of special targets arranged in a grid (the reference grid), which can be placed close to a wall or in any other place of the room. The thermal camera is mounted on a pan-tilt unit to realize the monitoring process in an automatic way. The system processes the thermal images in real-time and autonomously controls the pan-tilt unit. A fast automatic learning procedure enables to recognize the special targets on the grid also in challenging environments and in different environment conditions, while a Particle Filter is used to update the state of the system (i.e., position of the intersection point between the optical axis of the camera and the planar surface of the grid).

The system is able to perform a reliable global localization of the position of the thermal camera. During the scanning of the wall surfaces, a set of positions are automatically and sequentially reached by the moving IR camera: for each position a thermal image is recorded. Images are hence rectified in order to obtain a more accurate temperature sampling. We successfully tested our system in several challenging environments.

1 Introduction

Instruments commonly used to monitor environmental conditions can perform punctual measurements. Most of the times, they are not appropriate for creating a map of the distribution of the values of the parameters of interest (e.g., temperature, light, humidity, air speed) within a building. It is totally impractical to move a measuring head comprising a temperature sensor, anemometer, hygrometer every 10 cm in a given room. Moreover, variations of interest are sometimes below the range of common instrumentation. For instance, low limit for ordinary anemometers is about $0.5 \, ms^{-1}$ while natural convection occurring in buildings could be even smaller than $0.1 \, ms^{-1}$.

Another issue is the non-destructive moisture detection, because moisture is the main reason for building damages. Currently, the moisture survey of buildings, performed using equipment such as moisture meters, is slow and prohibits the monitoring of hard-to-reach locations. An infrared thermal imaging camera helps to monitor moisture by imaging the different temperatures of wet versus dry building materials, making it a useful tool for a building moisture survey. Non-destructive testing of water intrusion can be performed using the IR camera in modern buildings, but could fail when high relative humidity and low temperature decrease the evaporation process. More important, the relation between moisture and temperature is much more complex as expected and a much deeper knowledge is needed [12].

Figure 1: The IR camera in front of the reference grid used to monitor the environmental conditions of the Masino castle (Italy) with the proposed approach.
The study of the indoor thermal and humidity conditions is a problem addressed by many authors [4]. The distribution of the surface temperature is of course very important, but the correct approach is to measure the boundary air conditions, at the same time. It is a matter of fact that the main limiting factor is the difficulty to analyze the environmental conditions using a suitable time and space scale. In fact, for a detailed study the density of measurement required is much more extensive than the ones practically available using ordinary techniques.

The paper illustrates a new approach to solve such a problem using IR thermography (patent pending [7]). Basically, the new deal is to measure the surface and air thermodynamic status starting from temperature measurements on special set of targets (Fig. 3(b)), distributed on the inspected volume (e.g., Fig. 1) in order to compose a so called reference grid (Fig. 1 and 2). The scanning of the wall surfaces and targets by the IR camera enables to obtain a high-resolution map of the physical phenomena distribution [1]. Finally, the local conditions are recovered by means of an automatic processing, based on a robust mathematical model.

1.1 System overview

The proposed system performs the scanning of the surface of interest in an automatic way. It detects the targets in the right order, even if they are not in a regular pattern, in order to follow the geometry of a complex building (Fig. 1). The system detects a set of putative targets inside every thermal image using an efficient blob-detection technique. Every candidate is therefore classified as inlier or outlier using a trained Adaptive Boosting (AdaBoost) machine learning algorithm [6]. The metrical positions where the targets are located are then recovered in real-time. We propose to cope such a task exploiting the relationship between the IR camera image plane and the planar surface of the reference grid (Fig. 2). Our approach is to represent the state of the system (i.e., the position of the intersection point between the optical axis of the camera and the planar surface of the reference grid) in a probabilistic fashion through a posterior density function (PDF). The PDF can be estimated recursively over time from the incoming observations (thermal images) and actions (pan-tilt movement) using a Recursive Bayes Filter [13] that exploits the Markov assumption of the stochastic process. The action model represents the probability density of the state at time $t$ given the state at time $t-1$ and the last action (pan-tilt movement) performed. The observation model represents the probability density of the observation at time $t$ given the state at the same time. To solve the Bayes Filter we use a Particle Filter, that is an approximated solution of the Bayes Filter. Particle Filters are well suited to cope with multi-modal PDF, as in our case, due to the strong spatial symmetry of the reference grid.

Before starting the scanning process, the system is able to perform a reliable global localization of the position of the thermal camera.

During the scanning of the wall surfaces, a set of special positions are automatically and sequentially reached by the moving IR camera: for each position a thermal image is recorded. These special positions represent the centroids of the neighbour targets. Images taken in such locations usually contain four or more targets, enabling the estimation of the homography that allows an accurate mapping between image points and world points. Given the homography, we rectify the image (i.e., remove the distortion due to perspective projection) in order to obtain a more accurate temperature sampling.

Finally, a resampling process gives a very detailed spatial distribution of the air and surface thermal hygrometric conditions. Of course, the arrangement of different thermograms composing a mosaic of the entire wall surface underlies the hypothesis that no significant changes happened during the scanning time. Therefore, the faster the scanning, the better it is.

2 Hardware Setup

The prototype is made of four main hardware components: 1) the sensor device (the IR camera); 2) the pan-tilt camera mount; 3) the controlling PC and 4) the reference grid.

Figure 2: The hardware setup of the system: during the scanning process, the system automatically detects and tracks the positions of the targets that compose the reference grid. The red dashed line represents a possible scanning path.
2.1 Thermal camera

The chosen thermal camera is a FLIR A320 (Fig. 3(a), equipped with the standard lens (Field of View $25^\circ \times 19^\circ$; focal distance: 18 mm; $F$ number 1.3; spatial resolution: 1.36 mrad). The detector is an uncooled microbolometer focal plane array of $320 \times 240$ pixels, working in the $7.5 - 13 \mu m$ spectral band, with 70 mK of thermal resolution at room temperature. The frame rate is 9 Hz. This device is mainly conceived for automatic industrial process control, due to dedicated features and advanced interconnection.

2.2 Pan-tilt camera mount

The position of the thermal camera is controlled by a simple and cheap pan-tilt motor with a constant velocity of $6^\circ$ degrees per second for the horizontal movement (pan) and $3^\circ$ degrees for the vertical movement (tilt). The device is controlled by the PC via a RS232 serial connection and it doesn’t provide any feedback information (e.g., the odometry).

2.3 Controlling PC

The thermal images acquisition and the camera motion control are performed in real-time using a 2GHz Core 2 Linux-based laptop. The target positions are tracked and matched with a reference grid (Fig. 2, see below). The camera position is continually changed in agreement with a given sampling sequence. For every cluster of targets, a set of parameters are sampled.

2.4 Reference grid

The innovative features of the method are mainly based on the extraction of useful radiative data by means of a set of special targets (Fig. 2 and 3(b), patent pending [7]). A light metallic frame is placed in the field of view of the IR camera holding on the targets. This reference grid serves for the following purposes:

- makes easier the identification of the running surface patch at the particular time;
- allows image registration and correct geometric scaling;
- allows a precise temperature compensation for the reflected thermal energy;
- over the targets the IR camera is able to record particular temperature values needed for the measuring of the air temperature map; also the air speed parallel to the surface is mapped on each target point;

Furthermore, a reference passive device is used to calibrate the temperature values measured by thermography and to measure the relative humidity of the air [11, 2].

A precise thermographic reading of an object that is not a black-body, requires at first the knowledge of the reflected IR radiation by the inspected surface and the surface emissivity [10]. A practical way to achieve such a data is to use a diffusive-reflective material placed close to the surface. Any target supported by the grid is patched with multiple squared surfaces where these informative radiative fluxes can be detected by the thermal camera and used for the heat flux estimation. One of these surface is composed by a wet lint ant it will be exploited in the target tracking process.

3 Targets Detection

In order to automatically detect the targets inside the environment, we focus on the wet lint that is embodied in each target (Fig. 3(b)). In fact, such a slice appears usually well-defined inside the thermal images (e.g., Fig. 4(a)). Our process starts detecting a set of putative targets inside every thermal image using a blob-detection algorithm. Every candidate is therefore classified as inlier (real target) or outlier (false positive) using a trained Adaptive Boosting (AdaBoost) machine learning algorithm.

3.1 Selecting putative targets

Given the known distance of the IR camera from the grid, we can approximately fix the size of the targets in the thermal images, and the radius (say $\sigma$) of its squared surfaces as well. The surface composed by the wet lint usually appears roughly as a uniform and well-defined squared region. Therefore, we start the process looking into the image for all the blob structures with approximately radius $\sigma$. As blob detector, we use the Difference of Gaussians ($DoG$) operator, that is a computational efficient approximation of the Laplacian of the Gaussian ($LoG$) filter, one
of the most common blob detectors.

![Image of blob detection process](image)

**Figure 4:** (a) The input gray-level thermal image: two targets are visible in the left side of the image; (b) The Difference of Gaussians (DoG) filter applied to the input image; (c) The adaptive thresholding operator applied to the DoG image (b): every white blob represents a putative target; (d) The putative targets detected in (c) are classified as inlier or outlier based on surrounding pixels in (a) using the AdaBoost machine learning algorithm.

We can define \( l(x, y, \sigma) \) as the convolution of the original thermal image \( f(x, y) \) with a bivariate Gaussian kernel \( g(x, y, \sigma) \) with zero-mean and diagonal covariance matrix with all non-zero entries equals to \( \sigma^2 \):

\[
l(x, y, \sigma) = g(x, y, \sigma) * f(x, y)
\]

The Difference of Gaussians is therefore defined as:

\[
\text{DoG}(x, y; \sigma) = l(x, y; \sigma + \Delta \sigma) - l(x, y; \sigma - \Delta \sigma)
\]

where \( \Delta \sigma \) is chosen in order to well approximate the LoG operator. The application of the DoG filter results in strong positive responses for dark blobs (e.g., the wet lints of the targets) of radius \( \sigma \) (Fig. 4(b)).

Finally, we need to select a set of connected dark regions that represent our putative targets. We transform the DoG grayscale image to a binary image \( th(x, y) \) using an adaptive thresholding operator [9], according to the following equation:

\[
\text{th}(x, y) = \begin{cases} 
1 & \text{if } \text{DoG}(x, y; \sigma) > T(x, y) \\
0 & \text{otherwise}
\end{cases}
\]

where \( T(x, y) \) is a threshold calculated individually for each pixel as a gaussian weighted sum of pixels neighborhood \((b \times b)\), summed by a positive parameter \(s\). We typically use \(b = 25\) while \(s\) depends on the contrast of the wet lints inside the thermal images. An example of the application of the adaptive thresholding operator is depicted in Fig. 4(c).

For each detected blobs (i.e., connected dark regions), it is hence computed the centroid, the area in pixels and the size ratio \((\text{width}/\text{length})\) of its bounding box.

### 3.2 Classify targets

As we have seen, every detected blob represents a putative target: we turn now to the problem of the classification of those regions in real (inlier) and fake (outlier) targets. This is a 2-class categorical classification problem: we use a method similar to the Viola-Jones object detection technique [14], based on the AdaBoost (Adaptive Boosting, [6]) learning algorithm applied to a set of Haar-like features efficiently extracted from the input thermal images.

#### 3.2.1 Feature extraction

As most of the classification algorithms, also AdaBoost needs a discrete representation (i.e., a features vector) of the object that should be classified.

![Image of Haar-like features](image)

**Figure 5:** The four Haar-like features used in the classification: the white regions is interpreted as “add that area”, while the black regions as “subtract that area”.

For each putative target, we consider in the input gray-level image (Fig. 4(a)) a squared region of fixed size that should cover all the surface of the viewed target (e.g., the black and white boxes depicted in Fig. 4(d)), and not only the wet lint surface (see Fig. 3(b)).

Similarly to [14], we compute in that region a small number (four in our case) of Haar-like visual features (Fig. 5): the white regions is interpreted as “add that area”, while the black regions as “subtract that area”. The sums and differences of the pixel values over the squared regions are calculated efficiently using integral images [5]. We include in our features vector also the two values representing the area in pixels and the size ratio \((\text{width}/\text{length})\) computed during the selecting of the putative targets (Sec. 3.1).

#### 3.2.2 Classification using boosting

Boosting is a classification method that works by sequentially applying a simple classification algorithm (in most cases, a decision tree) to re-weighted versions of the training data, and then taking a weighted majority vote of the sequence of classifiers thus produced. Boosting is a supervised learning technique, that means it generates a function
that maps inputs (the features vector) to a class label (inlier or outlier in our case) given a set of training data. This dataset is composed of a number of pairs of input features vector and the desired class label.

AdaBoost (Adaptive Boosting) is an instance of the boosting machine learning algorithm that try to modify the sequence of the simple classifiers giving higher weight to training samples that are currently misclassified [6]. In our case the training stage is performed off-line before the scanning process.

![Figure 6](image.png)

Figure 6: The graphical user interface (GUI) of the application developed for the training stage of the AdaBoost algorithm. The user select each putative targets and assign to it a label (green boxes are inlier, gray boxes are outlier).

We develop a specific graphical tool for this task (Fig. 6). The user chooses a set of representative thermal images: for each, the application automatically select the putative targets (Sec. 3.1). Using a point-and-click strategy, the user decides the class label of the putative targets: inlier (green boxes in Fig. 6) or outlier (gray boxes in Fig. 6). The collected set of input feature vectors and the corresponding class labels (inlier/outlier) are therefore used to train the AdaBoost algorithm: in other words, the algorithm learn the functional relationship $F : y = F(x)$ between input feature vectors $x$ and the output labels $y$. This relationship will be used during the real classification stage to predict the class label of a putative target given its incoming feature vector.

4 Automatic Monitoring Process

The presented monitoring process involves an accurate computation for each location of some fundamental parameters as the air temperature and speed, the relative humidity and the wall temperature. Those thermodynamic quantities are evaluated given the measures extracted from the recorded thermal images during the scanning process. On the other hand, the IR camera focus only a partial fields of view (here’s called Partial Field-of-View, PFOV) of the desired area. In order to automatically perform the whole monitoring process, it is therefore essential to accurately track the location of every PFOV inside the area of interest during a complete scan.

Given the 2D metrical map of the reference grid (i.e., the location of every target) and a thermal image with the projection of at least four known targets (i.e., we know which are the projected target), it is possible to compute the homography $H$ between image points and grid points solving a linear system [8] (Fig. 7).

![Figure 7](image.png)

Figure 7: The IR camera looking at a plane with four point correspondences ($P_0$, $P_1$, $P_2$, $P_3$) needed to determine $H$. During the scanning process, we track the position of the intersection point between the optical axis of the camera ($\hat{z}_{cam}$) and the planar surface of the reference grid [source: www.wikipedia.com].

Given the homography $H$, it is possible to map every image point $p_i$ to a world point $P_i$ inside the surface of the reference grid as:

$P_i = H p_i$  \hspace{1cm} (4)

As we have seen, the computation of the homographies involves the knowledge of which target is actually inside the PFOV (i.e., which target is actually projected inside the thermal image). We therefore propose to estimate the position of the intersection point between the optical axis of the camera ($\hat{z}_{cam}$) and the planar surface of the reference grid: given the knowledge of this location (the state of the system), the four targets correspondences and hence the corresponding homography can be easily recovered.

It is important to note that no geometric calibration of the IR camera is needed in our system: it is only assumed that the optical axis of the camera $\hat{z}_{cam}$ intersect the image plane in a point fairly close to the image center.

4.1 Target tracking using a Particle Filter

Our approach is to represent the state of the system in a probabilistic fashion through a posterior density function
The key idea of the Particle Filter is to represent the posterior density of the state at time \( t \) given the state at time \( t-1 \) and the last action \( u_t \) (pan-tilt movement).

The observation model represents the probability density of the observation at time \( t \) given the state at the same time.

To solve the Bayes Filter (Eq. 5) we use a Particle Filter. The particles are finally resampled based on their importance-factor [13], i.e., we construct a new population of particles by selecting particles from the actual population with probability proportional to the weight of each particle. Some initial particles may be forgotten and some may be duplicated.

4.2 Initialization and scanning

At the beginning of the process, the particles are uniformly distributed in the whole working area (Fig. 8(a)). A sequence of fixed movements is performed in order to estimate the real position inside the grid: all particles that fall outside the working area obtain an importance-factor equals to zero, then they will be discarded during the resampling process. At the end of the initialization step, all particle are condensed around the real position (Fig. 8(f)). During the scanning of the wall surfaces, a set of special positions are automatically and sequentially reached by the moving IR camera: for each position a thermal image is recorded. These special positions (magenta boxes in Fig. 8) represent the centroids of the neighbour targets. Images taken in such location usually contain four or more targets, enabling the estimation of the homography that allows an accurate mapping between image points and world points. Given the homography, we rectify the image (i.e., remove the distortion due to perspective projection) in order to obtain a more accurate temperature sampling.
Figure 8: A typical localization sequence. The green boxes represent the position of the targets inside the reference grid, while the magenta boxes represent the special positions that should be reached during the scanning process. (a) The particles are uniformly distributed in the whole working area; (b), (c), (d), (e) Due to the IR camera movements, the particles tend to come close to real position; (f) All the particle are condensed around the real position.

5 Experiments

We implemented our monitoring system in C++ using the efficient OpenCV image processing library\(^1\) [3]: in Fig. 9 is depicted the graphical user interface of the application. The whole process runs in real-time with a frame-rate of 9 \(fps\) on a 2GHz Core 2 Linux-based laptop.

Figure 9: The graphical user interface (GUI) of the developed application. On the left: the actual grabbed IR-image; On the right: the reference grid and the particles representing the probability distribution of the position of the IR camera’s optical axis inside the grid.

We successfully tested our system in several environments with different distributions in the surface temperature (e.g., Fig. 1 shows the technique of measurement at work during the experiments inside the Masino castle in Italy). During the initialization step, the real position of the intersection point between the optical axis and the planar surface of the grid has been correctly estimated in all the experiments (Fig. 8(f)).

Figure 10: The condensation risk studied by the new method inside the Masino castle (Italy).

After the initialization, the complete scanning procedures has been completed correctly in all cases, following the expected scanning path and grabbing the thermal images in the correct positions. This results have been reached after a correct learning procedure that enables the system to well recognize the targets inside the grid.

Figure 10 shows a few cross sections that has been measured according to the described method in Masino castle in Italy. The maps represent the air temperature in the range 3.5 \(\div\) 5.5\(^°\)C, as seen in Fig. 11 (a). Similar maps for the air speed distribution in the range 0 \(\div\) 0.5\(ms^{-1}\) has

\(^1\)http://opencv.willowgarage.com
been collected (see Fig. 11 (b)). All results are geometrically corrected, with horizontal and vertical coordinates expressed in $cm$.

Figure 11: (a) Air temperature in the range $3.5 \div 5.5^\circ C$, with geometry corrected and space scale in $cm$; (b) Air speed distribution in the range $0 \div 0.5 ms^{-1}$ (space scale in $cm$).

6 Conclusions

IR Thermography is an effective non-destructive method for the evaluation of the indoor thermal-hygrometric conditions and the materials decay. Using the proposed grid of targets, the moisture damage could be detected in a much more reliable and non-destructive way. High thermal and spatial resolution are achieved allowing to follow the real physical process. The proposed robotized system allows an automatic tracking of the grid targets and a reliable and accurate mapping between image points and world points that improves significantly the effectiveness of the method. Indeed, with a fully automatic scanning the presence of humans inside the monitored environment is not necessary, avoiding perturbations of the microclimate natural status. Future works deal with the improvement of the target detection strategy in order to make it less sensitive to the quality of the learning procedure.

References


