

# USING DIRECTIONAL STATISTICS TO LEARN CAST SHADOWS FROM A MULTI-SPECTRAL LIGHT SOURCES PHYSICAL MODEL

*Rui Caseiro, João F. Henriques and Jorge Batista*

Institute of Systems and Robotics, DEEC-FCTUC, University of Coimbra, Portugal

## ABSTRACT

In this paper is proposed a novel statistical learning approach, to identify cast shadows, and model their generation. We exploit the theoretically well-founded directional statistics field, in order to formulate the generation of cast shadows as a Mixture of Von Mises-Fisher distributions (MovMF) on the unit sphere. This formulation is based on a bi-illuminant physical model of cast shadows, where no prior assumptions of the spectral power distribution (SPD) of the direct light sources and ambient illumination in the scene are made. Founded on a rigorous directional statistics approach, this parametric framework is capable of modelling the shaded surface behavior in complex illumination scenes and meet real time requirements. This better model discriminating cast shadows provides a more compact representation, and achieve better accuracy, with less data and much less computation time, compared with non-parametric models previously proposed. Theoretic analysis and experimental evaluations demonstrate the effectiveness of the proposed framework.

## 1. INTRODUCTION

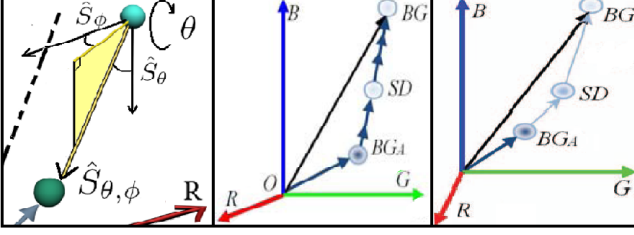
In vision-based applications, such as video surveillance, cast shadows are considered a major concern for foreground detection algorithms. They pose important problems, such as shape and color properties distortion of the objects, inducing silhouette distortions and object merging. When a foreground object casts a shadow on a background surface, the light sources are partially or entirely blocked, reducing the total energy incident, hence it is induced a variation of its appearance. Shadow points are expected to have similar chromaticity values but lower luminance. However texture characteristic remains unchanged since shadows do not alter the surfaces. This variation is one of the main properties used in the literature to detect cast shadows. It is dependent of the scene composition, such as the presence of other light sources and the reflectivity properties of other scene objects. There have been many proposed methods to detect cast shadows assuming that the value of the surface under cast shadows will be linearly attenuated from the background value, and thus fall on the line between the background value and the origin of the color space (linear model). In [1] the brightness component is separated from the chromaticity in RGB space defining brightness and chromaticity distortion. However, one main drawback of these methods is that require to set parameters for different scenes and can not handle complex and time-varying lighting conditions of the typical real scenarios. The statistical prevalence of cast shadows had been exploited to learn the surface appearance under cast shadows in [2],[3],[4]. Methods that assume the linear model may falsely label pixels as cast shadows when foreground objects have chromaticity values similar to that of the background. Furthermore, in some scenarios the background value under a cast shadow may not necessary

attenuate linearly along the line between the value of the corresponding background and the color space origin. This situation occurs because the light sources may consist of direct sunlight, diffused light scattered by the sky, and other colored light from nearby surfaces with different SPD. In outdoor scenes when blocking a surface from direct sunlight, the light scattered by the sky has a spectrum which differs from that of the sun. In order to overcome the linear model problems, in [5] was proposed a model in which the illumination is approximated by an ambient illumination term and a punctual light sources term. Unlike linear model case, these two terms can have a different SPD. It is only assumed that all punctual light sources have a common SPD with different power factor and the ambient illumination is constant over lit and shaded regions. Hence, the ambient illumination term determines the direction in color space along which the shaded background values can be found. Nonparametric density estimation was used to model surface variation under cast shadows. In [6] was considered that the shadow model presented in [5] may fail to deal with environments with complex illumination conditions. Thus, their work release the prior assumption that punctual light sources have the same profile with only different power factor and consider that ambient illumination may be not constant, due slightly changes caused for example by foreground objects circulation. They learn and update the shadow model parameters using a Gaussian mixture model (GMM) over time.

In this paper we follow the physics-based approach proposed in [6], but we introduce a novel method to identify cast shadows, modeling parametrically the shadow direction ( $\hat{S}$ ).  $\hat{S}$  is a unitary vector in a coordinate system centered on  $\mathbf{BG}$  (Fig.1), its direction is represented in spherical coordinates. Since we are dealing with data over the sphere,  $\mathbb{S}^2$ , we wish to express this data using distributions over this domain. A natural model for directional data is provided by the von Mises-Fisher distribution (vMF) on the unit hypersphere. We propose model this complex data using a mixture model tailored for directional data distributed on the surface of a unit hypersphere, consisting of a MovMF. This parametric model is introduced based on the idea of modeling shadow direction as a mixture of directional statistics distributions which provides a statistical interpretation of the shadow direction. Notice that, such decomposition cannot be achieved with non-parametric models. Our model is compact in that it requires very few parameters to represent complicated geometries. In particular, low-dimensional parametric models of shadow direction play vital roles since they enable the formulation of such inverse vision problems as parameter estimation.

## 2. DIRECTIONAL STATISTICS

Directional statistical techniques take into consideration the structure and nature of the sample spaces and develop probability models based on which angular data can be described. To analyze statistically directional data, the vMF [7] [8] has been developed by the analogy with the Gaussian distribution. More general and complex



**Fig. 1: Left :** Shadow Direction  $\hat{S}$  representation. **Middle:** Physical model proposed in [5]. **Right :** Physical model proposed in [6].

models over spherical domains, e.g., Fisher-Binghamor [7], may not be viable for real cases since the parameter estimation problem is significantly more difficult, and require substantially more training data. Unlike these models, the vMF has fewer variables and involves constraints which are more simpler to satisfy. The vMF is particularly useful for statistical inference of data that is inherently directional in nature and has been adopted in directional statistics because it results much more suitable to treat periodic variables. Gaussians are often used [6] since they are well known and suitable statistical treatment has been deeply analyzed. However, in the case of angular variables, Gaussians can be affected by errors due to the discontinuity in the origin, thus are inadequate for characterizing such data. Treating periodic variable by setting a value as origin and then applying traditional gaussian will bring to results that were strongly dependent on the arbitrary choice of the origin. This limitation does not affect vMF because it is independent on the origin. In fact, vMF yields many of the key properties for statistical inference that the normal distribution has for linear data. We tackle the challenging problem of deriving a low-dimensional parametric model that can achieve accuracy comparable to non-parametric models [5]. Since non-parametric models are essentially raw measurements, they undoubtedly have strong advantages. However, when using such non-parametric models for solving some inverse problems in computer vision, we are cursed by the high-dimensionality of data. Furthermore, the accuracy of non-parametric models essentially depend on the sampling of the data and thus necessitates very dense measurements to achieve certain accuracy. The statistical model presented lays the foundation for deriving canonical probabilistic formulations, for this physics-based vision problem, in order to estimate a correct parametrization for it. The vMF is a probability distribution on the hypersphere  $\mathbb{S}^{d-1}$ , embedded in  $\mathbb{R}^d$  for directional data distributed unimodally with rotational symmetry. In its most general form, the probability density function of the vMF distribution of a  $d$ -dimensional unit random vector  $\mathbf{x}$ , generalized to a hypersphere  $\mathbb{S}^{d-1}$  (i.e.,  $\mathbf{x} \in \mathbb{R}^d$  and  $\|\mathbf{x}\| = 1$ , or equivalently  $\mathbf{x} \in \mathbb{S}^{d-1}$ ) is given by

$$\left[ p(\mathbf{x}|\mu, \kappa) = c_d(\kappa) e^{\kappa \mu^T \mathbf{x}} \right] \left[ c_d(\kappa) = \frac{\kappa^{d/2-1}}{(2\pi)^{d/2} I_{d/2-1}(\kappa)} \right] \quad (1)$$

where  $\|\mu\| = 1$ ,  $\kappa \geq 0$  and  $d \geq 2$ . Thus,  $\mathbf{x}, \mu \in \mathbb{S}^{d-1}$  and  $c_d(\kappa)$  is the normalizing constant.  $I_r(\cdot)$  is the  $r^{\text{th}}$  order modified Bessel function of the first kind. The density  $p(\mathbf{x}|\mu, \kappa)$  is parameterized by the mean direction  $\mu$ , and the concentration parameter,  $\kappa$  (analogous to the precision or inverse of the variance in Gaussian cause) so-called because it characterizes how strongly the unit vectors drawn according to  $p(\mathbf{x}|\mu, \kappa)$  are concentrated about the mean direction  $\mu$ , indicates the degree of directional dispersion. In particular when  $\kappa = 0$ ,  $p(\mathbf{x}|\mu, \kappa)$  reduces to the uniform density on  $\mathbb{S}^{d-1}$  (isotropic scattering) and as  $\kappa \rightarrow \infty$ ,  $p(\mathbf{x}|\mu, \kappa)$  tends to a point density (scattering becomes extremely non-isotropic). The ordinary sphere oc-

curs when  $d = 3$ , which is of our immediate interest in shadow direction problem. In this case the probability density function of the vMF distribution has the form

$$p(\mathbf{x}|\mu, \kappa) = \frac{\kappa}{4\pi \sinh(\kappa)} e^{\kappa \mu^T \mathbf{x}} \quad (2)$$

Considering a MovMF as the underlying generative model for directional data, the probability density function of the MovMF model is given by  $p(\mathbf{x}|\Theta) = \sum_{j=1}^K \alpha_j p(\mathbf{x}|\theta_j)$  where  $\Theta = \{\alpha_1, \dots, \alpha_K, \theta_1, \dots, \theta_K\}$  where  $\alpha_j$  is the weight (prior probability) of the  $j^{\text{th}}$  vMF state in the mixture model, with  $\sum_{j=1}^K \alpha_j = 1$ ,  $\alpha_j \geq 0$ , and  $p(\mathbf{x}|\theta_j)$  is a single vMF with parameters  $\theta_j = (\mu_j, \kappa_j)$ .

The clustering of data lying on the surface of a hypersphere is posed as a maximum likelihood estimation problem, which requires the definition of a complete procedure to derive the mixture parameters. The Expectation Maximization (EM) is a general technique for finding maximum likelihood estimators in latent variable models, in cases where analytic solutions are difficult or impossible. It has been widely used to estimate the mixture parameters due to its simplicity and numerical stability. Banerjee in [8] proposed two variants of the EM algorithm and appropriate approximations to estimate the mixture parameters. These two variants are centered on soft and hard-assignment schemes (respectively called soft-movMF and hard-movMF). The soft-movMF algorithm, assigns soft (or probabilistic) labels to each point given by the posterior probabilities (3). In soft-movMF case the mixture parameters are estimated exactly following the Banerjee's derivations given by

$$p(j|\mathbf{x}, \Theta) = \frac{\alpha_j p(\mathbf{x}|\theta_j)}{\sum_{l=1}^k \alpha_l p_l(\mathbf{x}|\theta_l)} \quad (3)$$

$$\left[ \alpha_j = \frac{1}{n} \sum_{i=1}^n p(j|\mathbf{x}_i, \Theta) \right] \quad \left[ \mu_j = \frac{\sum_{i=1}^n \mathbf{x}_i p(j|\mathbf{x}_i, \Theta)}{\sum_{i=1}^n p(j|\mathbf{x}_i, \Theta)} \right] \quad (4)$$

$$\left[ \bar{r} = \frac{\|\mu_j\|}{n \alpha_j} \right] \quad \left[ \mu_j = \frac{\mu_j}{\|\mu_j\|} \right] \quad \left[ \kappa_j = \frac{\bar{r} d - \bar{r}^3}{1 - \bar{r}^2} \right] \quad (5)$$

In this paper we use the hard-movMF. The hard assignment scheme is a winner-takes-all strategy. That is, in E-step, a data point is assigned to one and only one cluster which has the highest posterior probability given by (3), i.e. it is performed a disjoint  $k$ -partitioning of data based on the posteriors. The crucial difference in this case is that the posterior probabilities  $p(j|\mathbf{x}, \Theta)$  are allowed to take only binary values (0/1). Numerical estimation of  $\kappa$  is non-trivial, since it involves functional inversion of ratios of Bessel functions. We here use a asymptotic approximation, proposed in [8].

### 3. ESTIMATION OF MODEL PARAMETERS

The first step in our approach is to estimate the shadow direction ( $\hat{S}$ ). However, to accurately capture this direction, samples obtained from background modeling process must be filtered out. We apply a weak classifier as a pre-filter to evaluate every moving pixels to select possible and filter out some impossible shadow points. The classifier used [9] determine whether a pixel is possibly shadow by measuring the similarity of color and texture between foreground and background. This is done by estimating the Color Normalized Cross-Correlation (CNCC) between each segmented pixel  $\mathbf{I}_t$  and the corresponding background pixel  $\mathbf{B}_t$ . More precisely, a pixel is classified as shadow if its texture is correlated with the corresponding texture of  $\mathbf{B}_t$ . The main goal is to determine whether a pixel value fits with the property of cast shadows and not necessarily to detect shadows accurately. Once the shadow direction has been estimated, we use a decomposition similar to [5], with a illumination

attenuation  $\beta$  term and a chromaticity distortion  $\mathbf{CD}$  term. The illumination attenuation is obtained as in [5], but in our case the term  $\mathbf{CD}$  is computed probabilistically, using the MovMF models.

### 3.1. Cast Direction Estimation : ( $\hat{\mathbf{S}}$ )

After clustering the data lying on the surface of the sphere, the next step is to estimate for each pixel the shadow direction  $\hat{\mathbf{S}}$  associated to its background values, i.e. the direction, in color space, along which the value of a background surface will be found under cast shadows. The MovMF is based on the assumption that each pixel views shadows states more often than foreground ones, therefore the states with higher prior probabilities (weights) and higher concentrations would be considered as shadows. Based on this premise, the first  $\mathbf{B}$  states with higher weights and concentrations, whose combined priori probabilities  $\alpha_j$  is greater than a pre-defined weight threshold  $\mathbf{T}$ , ( $\mathbf{B} = \arg \min_b \sum_{j=1}^b \alpha_j > \mathbf{T}$ ), are considered as the representative models of shadows. For each sample  $\mathbf{x}_i$  we compute its direction  $\hat{\mathbf{s}}_i$  relative to  $\mathbf{BG}$ . The idea is to attempt to match the current observation  $\hat{\mathbf{s}}_i$  at a given pixel with one of the  $\mathbf{B}$  states. If the observed sample is inside of 99% confidence interval of one of these distributions, we consider the mean of this distribution as the cast shadow direction in color space on which background surface values under cast shadows are found. The chromaticity distortion  $\mathbf{CD}$  term is now calculated as the likelihood of the sample with respect to this distribution. Pixels that are outside of 99% confidence interval of all  $\mathbf{B}$  distributions are automatically classified as foreground pixels ( $\mathbf{CD} = 0$ ) Numerical estimation of the confidence interval in vMF case is non-trivial, but several closed and fast solutions have been proposed. For exact equations, please refer to [7].

### 3.2. Illumination Attenuation Estimation : ( $\beta$ )

Even if samples are situated on the shadow direction, this does not mean that these samples represent cast shadows. If this decision is only based on this criterion would label objects with different shades of gray as shadow. To further improve the differentiating ability, we model parametrically the posterior distribution of lighting attenuation  $\beta$ , for both cast shadow  $P(z = \mathbf{SD}|\beta)$  and foreground  $P(z = \mathbf{FG}|\beta)$  hypothesis ( $\beta$  is computed as in [5]). The pixel-based illumination variation is modeled by a one dimensional GMM, separately for each one of these hypothesis, using an online K-means algorithm as in [10]. Note that, henceforth the indice  $z$  presented in the next equations/variables can take values  $z = \mathbf{FG}$  (foreground) or  $z = \mathbf{SD}$  (shadow). The first  $\mathbf{B}_z$  states of the GMM with higher weights and smaller variances in the mixture are considered as the representative models of the illumination variation. The index  $\mathbf{B}_z$  is determined by  $\mathbf{B}_z = \arg \min_b \sum_{j=1}^b \alpha_j^z > \mathbf{T}_z$ , where  $\mathbf{T}_z$  is a pre-defined threshold. The likelihood  $P(\beta|z)$  is generated as follows

$$P(\beta|z) = \frac{1}{\mathbf{W}^z} \sum_{j=1}^{\mathbf{B}_z} \alpha_j^z \mathcal{N}(\beta|\theta_j^z) \left\{ \begin{array}{l} \mathbf{r}_w = \mathbf{C}_{i,z} \left( \frac{1 - \mathbf{r}_0}{\sum_{l=1}^K \mathbf{c}_l} \right) + \mathbf{r}_0 \\ \mathbf{r}_g = \mathbf{C}_{i,z} \left( \frac{1 - \mathbf{r}_0}{\mathbf{c}_j} \right) + \mathbf{r}_0 \end{array} \right. \quad (6)$$

where  $\mathbf{W}^z = \sum_{j=1}^{\mathbf{B}_z} \alpha_j^z$  is a normalization constant.  $\mathbf{r}_w$ ,  $\mathbf{r}_g$  are the learning rates, respectively for the mixing weights and the Gaussian parameters (means and variances).  $\mathbf{c}_j$  is the number of matches of the  $j^{\text{th}}$  Gaussian state, and  $\mathbf{r}_0$  is a small constant ( $\approx 0.005$ ). An effective learning algorithm is used since the conventional GMM learning approach when applied to pixel-based models may suffer from slow learning due to insufficient training data. Since new samples may not appear at the same pixel in each frame, when foreground activities are rare. To address this problem, the GMM is

updated through a confidence-rated learning approach. The updating scheme follows the formulation of the combination of incremental EM learning and recursive filter as proposed in [11], which is significantly faster than conventional online updating. This method will follow the incremental EM learning, in the initial learning stage and approaches to recursive filter over time. Note that, since in our method, we model the illumination variation separately for each one of two hypothesis, unlike in [6] we not need use logistic regression similar to that in [11], to extract the likelihoods  $P(\beta|z)$  from GMM. The learning rates (6) are controlled by a confidence value equal to  $\mathbf{C}_{i,z} = P(\mathbf{FS}|\mathbf{x}) * \mathbf{CD} * \mathbf{CNCC}$  in ( $z = \mathbf{SD}$ ) case. We incorporate cross-correlation ( $\mathbf{CNCC}$ ) information for improving the shadow discriminative ability and the convergent rate. More weight is given to samples which are close to the cast shadow direction, are not background values and are correlated with background texture. Instead of perform a blind update, which treats each sample in the same way, this way permit weight each sample with different importance. If the model is updated using a potential shadow point, then this sample is considered relatively more important than others. The model needs not to obtain numerous samples to converge, but a few samples having high confidence value are sufficient. Observations with higher confidence values will converge faster than those with low ones. In ( $z = \mathbf{FG}$ ) case this confidence is given by  $\mathbf{C}_{i,z} = P(\mathbf{FS}|\mathbf{x}) * \mathbf{CD} * (1 - \mathbf{CNCC})$ . As a result, both illumination attenuation likelihoods are generated from samples that are either cast shadows or foreground samples sharing similar illuminance and color characteristics. The posteriors on shadow  $P(z = \mathbf{SD}|\beta)$  and foreground  $P(z = \mathbf{FG}|\beta)$  are calculated applying the Bayes theorem as follows:

$$P(z|\beta) = \frac{P(z)P(\beta|z)}{P(\mathbf{SD})P(\beta|\mathbf{SD}) + P(\mathbf{FG})P(\beta|\mathbf{FG})} \quad (7)$$

where  $P(\mathbf{FG})$  and  $P(\mathbf{SD})$  are the priors, computed by summing the sample confidences ( $P(z) = \sum_i \mathbf{C}_{i,z}$ ), and then normalized such that ( $P(\mathbf{FG}) + P(\mathbf{SD}) = 1$ ). These posteriors as restricted, such that  $P(\mathbf{SD}|\beta < 0) = 0$  and  $P(\mathbf{FG}|\beta < 0) = 1$ , i.e. samples brighter than the background cannot be cast shadow.

## 4. POSTERIOR PROBABILITIES

In this section, we combine the posteriors  $P(\mathbf{SD}|\beta)$  and  $P(\mathbf{FG}|\beta)$  with  $\mathbf{CNCC}$  and  $\mathbf{CD}$  information, in order to compute the posterior distributions  $P(\mathbf{SD}|\mathbf{x})$  and  $P(\mathbf{FG}|\mathbf{x})$ , respectively under shadow and foreground hypothesis. These posteriors are computed by decomposing respectively  $P(\mathbf{SD}|\mathbf{x})$  and  $P(\mathbf{FG}|\mathbf{x})$  over the ( $\mathbf{BG}, \mathbf{FS}$ ) domain ( $\mathbf{FS} = \mathbf{FG} \cup \mathbf{SD}$ ). They can then be used directly to segment cast shadow samples from non-background samples. Note that  $P(\mathbf{SD}|\mathbf{x}, \mathbf{BG}) = 0$  and  $P(\mathbf{FG}|\mathbf{x}, \mathbf{BG}) = 0$ . The probability that a pixel belongs to either the background  $P(\mathbf{BG}|\mathbf{x})$  or the foreground  $P(\mathbf{FS}|\mathbf{x})$ , are computed from a GMM [10], used to background modeling.

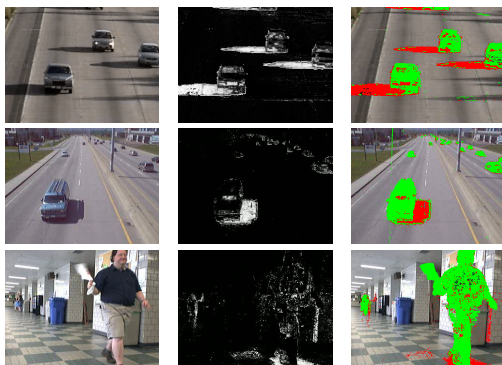
### 4.1. Cast Shadow Posterior : $P(\mathbf{SD}|\mathbf{x})$

The cast shadow posterior is given by  $P(\mathbf{SD}|\mathbf{x}) = P(\mathbf{SD}|\mathbf{x}, \mathbf{FS}) * P(\mathbf{FS}|\mathbf{x})$ . The term  $P(\mathbf{SD}|\mathbf{x}, \mathbf{FS})$  is decomposed into two parts: ( $\mathbf{D}$ ) and ( $\mathbf{ND}$ ), which stand for samples that are situated on the cast shadow direction or not, respectively. If a sample are outside of confidence interval of all MovMF states, then the probability of belonging to shadow equals to zero ( $P(\mathbf{S}|\mathbf{x}, \mathbf{ND}, \mathbf{FS}) = 0$ ). Therefore we have  $P(\mathbf{SD}|\mathbf{x}, \mathbf{FS}) = P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}) * P(\mathbf{D}|\mathbf{x}, \mathbf{FS})$ , where  $P(\mathbf{D}|\mathbf{x}, \mathbf{FS})$  is given by  $\mathbf{CD}$ . In order to calculate  $P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS})$ , we take into account the cross-correlation

information. Samples can be seen as belonging two categories: correlated (**C**) or not (**NC**). It follows that  $P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}) = P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{C}) * P(\mathbf{C}|\mathbf{x}, \mathbf{D}, \mathbf{FS}) + P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{NC}) * P(\mathbf{NC}|\mathbf{x}, \mathbf{D}, \mathbf{FS})$ . Since  $\beta$  can be seen as the sufficient statistics for  $\mathbf{x}$ , the term  $P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{C})$  is equal to  $P(\mathbf{SD}|\beta)$ . The value of  $P(\mathbf{SD}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{NC})$  is kept different of zero and small ( $= \kappa P(\mathbf{SD}|\beta)$ ) with ( $\kappa \leq 0.15$ ), in order to take into account the possibility that we could observe uncorrelated samples for cast shadow samples (noise). Finally,  $P(\mathbf{C}|\mathbf{x}, \mathbf{D}, \mathbf{FS}) = \mathbf{CNCC}$  and  $P(\mathbf{NC}|\mathbf{x}, \mathbf{D}, \mathbf{FS}) = (1 - \mathbf{CNCC})$ .

#### 4.2. Foreground Posterior : $P(\mathbf{FG}|\mathbf{x})$

The procedure to estimate the  $P(\mathbf{FG}|\mathbf{x})$  is similar to the  $P(\mathbf{SD}|\mathbf{x})$  case. If a sample are outside of confidence interval of all MovMF states, then the probability of belonging to foreground equals to one ( $P(\mathbf{FG}|\mathbf{x}, \mathbf{ND}, \mathbf{FS}) = 1$ ). The term  $P(\mathbf{FG}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{NC}) = P(\mathbf{FG}|\beta)$ , and the value of  $P(\mathbf{FG}|\mathbf{x}, \mathbf{D}, \mathbf{FS}, \mathbf{C})$  is kept different of zero and small ( $= \kappa P(\mathbf{FG}|\beta)$ ) with ( $\kappa \leq 0.15$ ), in order to take into account the possibility that we could observe correlated samples for foreground samples.



**Fig. 2:** **Left** : Original frame. **Middle**: Shadow Posterior  $P(\mathbf{SD}|\mathbf{x})$ . **Right** : Binary Classification - Shadow (red) - Foreground (green)

Methods	Highway I		Highway II		Hallway	
	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$	$\eta\%$	$\xi\%$
Proposed	75.63	84.15	76.13	80.67	72.85	89.94
PhGMM [6]	72.34	84.98	72.70	79.89	71.69	88.15
PhKernel [5]	70.50	84.40	68.40	71.20	72.40	86.70
LGf [2]	72.10	79.70	-	-	-	-
GMSM [3]	63.30	71.30	58.51	44.40	60.50	87.00

**Table 1:** Quantitative results

## 5. EXPERIMENTAL RESULTS

In order to analyze the effectiveness of the proposed method, we conduct several experiments on three video sequences, presented in previous literature, including both indoor and outdoor environments. The accuracy of our approach is compared quantitatively to other methods to shadow detection cited earlier when results are available. This evaluation use the most employed quantitative metrics utilized to evaluate the shadow detection performance, namely shadow detection ( $\eta$ ) and discrimination rate ( $\xi$ ), presented in [12]. Cast shadows induce a significant color shift, therefore breaking the shadow linear approximation. Since this color shift is modeled by our approach, we generate posterior distributions that are faithful to the scene. Our method can deal with the situation that shadows first appear in complex scenes and unknown illumination conditions as well as rare foreground activity. Through experimentation on benchmark data we clearly demonstrate the advantage of using a direc-

tional statistics based approach. The results were obtained by thresholding the shadow posterior and show that our model performs better than the parametric approach based on GMM [6]. Our method can achieve higher accuracy comparable to the state-of-the-art non-parametric model [5], specially when the sampling of measurements is sparse and with a much smaller footprint. According to the overall performance, our approach is more effective in describing background surface variation under cast shadows, compared to the other analysed methods. Qualitative and quantitative results presented validate the model we have introduced based on the physical properties of the light sources and surface behavior. The results also show that we can successfully learn the model parameters, i.e. the shadow direction and the illumination attenuation with respect to a background sample, under both shadow and foreground hypothesis. Note that the proposed approach is pixel-based, and the results presented here are raw data without any postprocessing. However, we stress that the posterior probabilities can be incorporated with a context model that incorporate spatial and temporal coherence using smoothness constraints, to improve the accuracy, and yield impressive results.

## 6. CONCLUSION

We introduced a novel method to identify cast shadows and model their generation using directional statistics. Understanding surface variation under shadows, requires that we understand how materials appear under realistic illumination conditions. Absent that understanding, shadow detection become more difficult since simple assumptions about how material colors behave under varying illumination create apparently random effects. We exploit a new physical model of cast shadows, free from prior assumptions of the SPD of illumination sources. Taking into account the structure and nature of the sample space we model parametrically the cast shadow direction. It can adapt to the illumination changes, particularly for the background under complex lighting conditions. This low-dimensional model requires very few parameters to represent complicated geometries. It achieves higher accuracy compared to non-parametric models, with a much smaller footprint and less time requirements.

## 7. REFERENCES

- [1] T. Horprasert, D. Harwood, and L. Davis, "Statistical approach real-time robust background subtraction shadow detection," in *ICCV*, 1999.
- [2] Zhou Liu, K. Huang, T. Tan, and Liangsheng Wang, "Cast shadow removal combining local and global features," *IEEE-CVPR*, 2007.
- [3] Nicolas Brisson and André Zaccarin, "Learning and removing cast shadows through a multidistribution approach," *IEEE-TPAMI*, 2007.
- [4] F. Porikli and J. Thornton, "Shadow flow: A recursive method to learn moving cast shadows," *IEEE-ICCV*, 2005.
- [5] Nicolas Brisson and André Zaccarin, "Kernel-based learning of cast shadows from a physical model of light sources and surfaces for low-level segmentation," *IEEE-CVPR*, 2008.
- [6] Chu-Song Chen Jia-Bin Huang, "A physical approach to moving cast shadow detection," *IEEE-ICASSP*, 2009.
- [7] K. V. Mardia and P. Jupp, *Directional Statistics*, John Wiley, 2000.
- [8] A. Banerjee, I. Dhillon, J. Ghosh, and S. Sra, "Clustering on the unit hypersphere using von mises-fisher distributions," *JMLR*, 2005.
- [9] D. Grest, J. Frahm, and R. Koch, "A color similarity measure for robust shadow removal in real time," *Vision Modeling Visualization*, 2003.
- [10] C. Stauffer and W. Grimson, "Adaptive background mixture models for real-time tracking," *IEEE-CVPR*, 1999.
- [11] D. S. Lee, "Effective gaussian mixture learning for video background subtraction," *IEEE-TPAMI*, 2005.
- [12] Andrea Prati, I. Mikic, Mohan Trivedi, and Rita Cucchiara, "Detecting moving shadows: Algorithms and evaluation," *IEEE-TPAMI*, 2003.