Fast Classification and Depth Estimation for Multispectral Single-Photon LiDAR Data

Amir Belmekki, Stephen McLaughlin, Abderrahim Halimi

School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh U.K.

HERIOT WATT UNIVERSITY

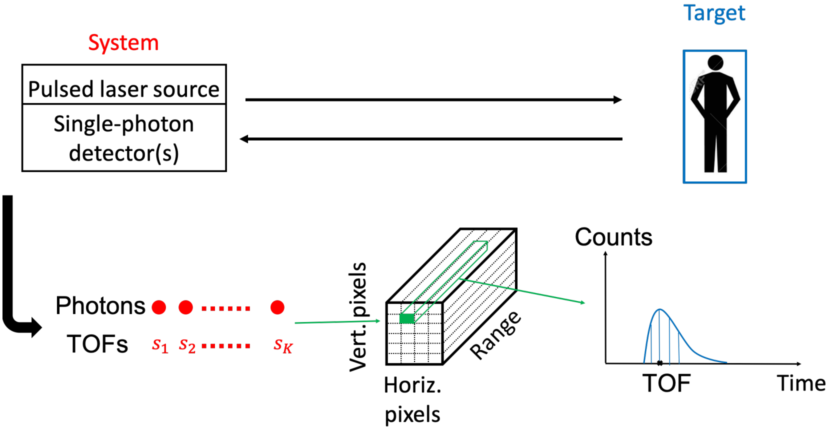
14 September 2021

Outline

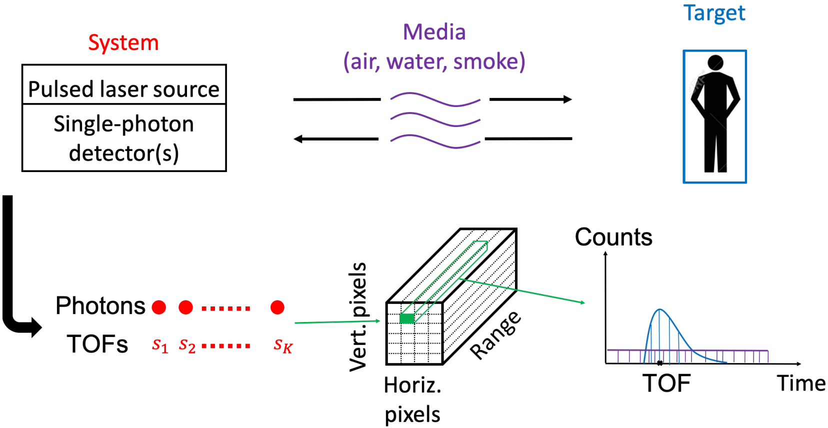
Plan

1. LiDAR principle and challenges
2. Proposed solution
3. Results
4. Conclusions and future work

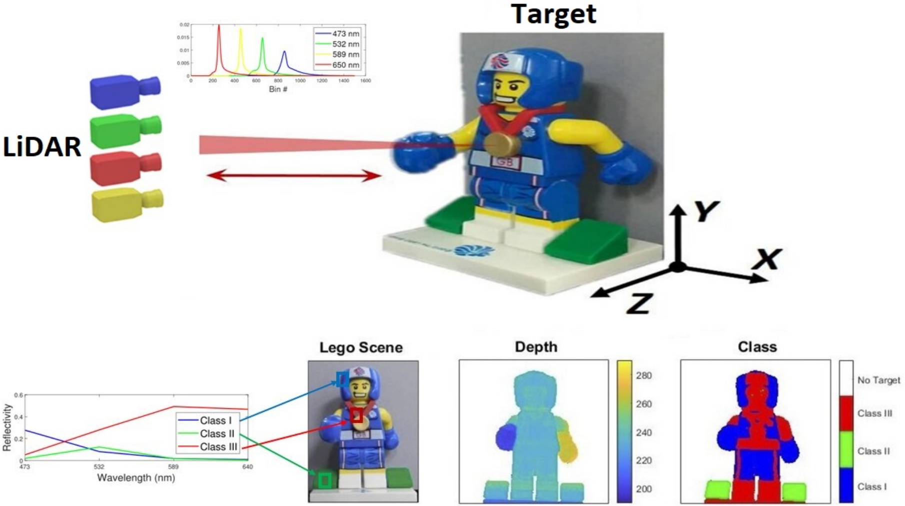
3D Lidar Imaging : principle



3D Lidar Imaging : Challenges



3D Lidar Imaging : Objectives[1]



A portion of this figure has been taken from [1] X. Ren, Y. Altmann, R. Tobin, A. Mccarthy, S. Mclaughlin, and G. S. Buller, "Wavelength-time coding for multispectral 3D imaging using single-photon LiDAR", Opt. Express, vol. 26, no. 23, Nov. 2018.

3D Lidar Imaging : Objectives

Fast segmentation, depth estimation and uncertainty quantification

- Spatial classification based on the known signatures
- Provide additional surface estimates (e.g. : depth estimate)
- Deliver uncertainty measure about the estimates
  $\rightarrow$ Possibility to perform a signature-based object detection.
  $\rightarrow$ Help acquire multidimensional data (i.e, video applications)

Plan

Solution Using a Bayesian/Regularization approach

Bayesian inference

- Estimate θ from observations y using the likelihood $f(y|\theta)$

Solution Using a Bayesian/Regularization approach

Bayesian inference

- Estimate θ from observations y using the likelihood $f(y|\theta)$
- The estimation of θ is an ill-posed problem
  - Noisy, corrupted or incomplete data

Solution Using a Bayesian/Regularization approach

Bayesian inference

- Estimate θ from observations y using the likelihood $f(y|\theta)$
- The estimation of θ is an ill-posed problem
  - Noisy, corrupted or incomplete data
- Introduction of a priori information $f(\theta)$ to regularize the problem $f(\theta|y) = f(y|\theta)f(\theta)/f(y)$

Solution Using a Bayesian/Regularization approach

$$f(\theta|y) = f(y|\theta)f(\theta)/f(y)$$

Exploit the posterior distribution

- Simulate the posterior distribution using stochastic simulation methods (e.g., MCMC)
  $\Longrightarrow$ computationally expensive
- Point estimators (e.g., MAP, MMSE) using optimization tools
  $\Longrightarrow$ often iterative, no uncertainty
- Carefully build a Bayesian model leading to analytical formulas of the estimates (and eventually their uncertainties)
  $\Longrightarrow$ fast implementation

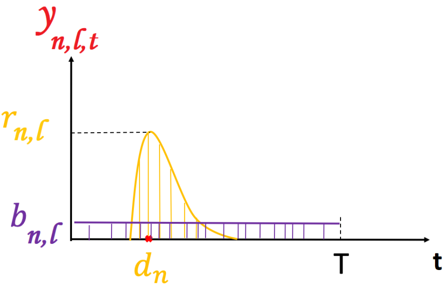
Observation model : Likelihood

$$y_{n,t,l}|r_{n,l}, d_n, b_{n,l} \sim \mathcal{P}[r_{n,l}\, g_l(t - d_n) + b_{n,l}],$$

Reformulated as

$$y_{n,t,l}|\omega_{n,l}, d_n, b_{n,l} \sim \mathcal{P}\{b_{n,l}\, [\omega_{n,l}\, T\, g_l(t - d_n) + 1]\}$$

where $\omega_{n,l} = \frac{r_{n,l}}{b_{n,l} T}$ is the signal-to-background ratio (SBR).



Observation model : Likelihood

$$y_{n,t,l}|r_{n,l}, d_n, b_{n,l} \sim \mathcal{P}[r_{n,l}\, g_l(t - d_n) + b_{n,l}],$$

Reformulated as

$$y_{n,t,l}|\omega_{n,l}, d_n, b_{n,l} \sim \mathcal{P}\{b_{n,l}\,[\omega_{n,l}\, T\, g_l(t - d_n) + 1]\}$$

where $\omega_{n,l} = \frac{r_{n,l}}{b_{n,l} T}$ is the signal-to-background ratio (SBR).

- Absence of target $r_{n,l} = \omega_{n,l} = 0$

  with $y_{n,t,l}|\omega_{n,l} = 0, d_n, b_{n,l} \sim \mathcal{P}(b_{n,l})$.
- Mixture of target and background counts $\omega_{n,l} > 0$
- Parameters : $\boldsymbol{\theta} = (\boldsymbol{\omega_n}, d_n)$, with $\boldsymbol{\omega_n} \in \mathbb{R}_+^L$ and $d_n \in [1, \cdots, T]$
- The posterior distribution $f(\boldsymbol{\theta}|y) = f(y|\boldsymbol{\theta})f(\boldsymbol{\theta})/f(y)$

Prior distributions

Gamma prior for the background $b_{n,l}$

$$P(b_{n,l}|\alpha_l^b, \beta_{n,l}^b) = \mathcal{G}(b_{n,l}, \alpha_l^b, \beta_l^b)$$

Spike-and-slab prior for the reflectivity $r_{n,l}$

$$P(r_{n,l}|u_n, \alpha_{k,l}^r, \beta_{k,l}^r) = \delta(u_n)\delta(r_{n,l}) + \sum_{k=1}^{K} \delta(u_n - k)\mathcal{G}(r_l, \alpha_{k,l}^r, \beta_{k,l}^r)$$

- $u_n = 0$ in absence of a target and $u_n = k$ when a target is present and its corresponding pixel has a class k
- $\alpha_{k,l}^r, \beta_{k,l}^r, \alpha_l^b, \beta_l^b$ are user fixed positive constants related to the known object spectral signatures.

Prior distributions

Joint distribution of the background $\boldsymbol{b_n}$ and the SBR $\boldsymbol{\omega_n}$

$$P(\boldsymbol{r_n}, \boldsymbol{b_n} | u_n, \phi, K) = P(\boldsymbol{r_n} | u_n, \phi, K) P(\boldsymbol{b_n} | \phi)$$

$$\boldsymbol{\omega_n} = \frac{\boldsymbol{r_n}}{\boldsymbol{b_n} T}$$

$$\downarrow$$

$$P(\boldsymbol{\omega_n}, \boldsymbol{b_n} | u_n, \phi, K) = \prod_{l=1}^{L} \Big[\delta(u_n) \delta(\omega_{n,l}) \mathcal{G}(b_{n,l}, \alpha_l^b, \beta_l^b) + \sum_{k=1}^{K} \delta(u_n - k) C_{k,l}(\omega_{n,l}) \mathcal{G}(b_{n,l}, \alpha_{l,k}^{\dagger}, \beta_{l,k}^{\dagger}(\omega_{n,l}) \Big]$$

Prior distributions

Uniform prior for the class label u_n

$$p(u_n = k) = \frac{1}{K+1} \quad ,\forall k \in \{0, ..., K\}$$

Uniform prior for the depth d_n

$$p(d_n = t) = \frac{1}{T} \quad ,\forall t \in \{1, ..., T\}$$

The likelihood and posterior distribution

The likelihood

$$p(\boldsymbol{Y}|\boldsymbol{\Omega}, \boldsymbol{B}, \boldsymbol{d}) = \prod_{t=1}^{T}\prod_{l=1}^{L} p(y_{n,t,l}|\omega_{n,l}, b_{n,l}, d_n)$$

where $\boldsymbol{d} = (d_1, \cdots, d_N)$ and $\boldsymbol{\Omega}$, $\boldsymbol{B}$ are two matrices gathering $\omega_{n,l}, \forall n, l$, and $b_{n,l}, \forall n, l$, respectively.

Posterior distribution

$$\mathrm{p}(\boldsymbol{\omega_n}, \boldsymbol{b_n}, d_n, u_n|\boldsymbol{Y_n}, \phi, K) \propto p(\boldsymbol{Y_n}|\boldsymbol{\omega_n}, \boldsymbol{b_n}, d_n, u_n)p(\boldsymbol{\omega_n}, \boldsymbol{b_n}|\phi, u_n, K)p(d_n)p(u_n)$$

where ϕ gathers all the hyperparameters of the model.

Decision rule

Depth estimation

$$\hat{d}_n = \underset{d}{\text{argmax}}\ p(d_n | \boldsymbol{Y_n})$$

Class estimation

$$H_{n,k} = \underset{k}{\text{argmax}}\ p(u_n = k \mid \boldsymbol{Y_n})$$

where

$$p(u_n|\boldsymbol{Y_n}) = \underbrace{\int \underbrace{\sum_{d_n=1}^{T} \int \overbrace{p(\boldsymbol{\omega_n}, \boldsymbol{b_n}, d_n, u_n|\boldsymbol{Y_n}) d\boldsymbol{b_n}}^{\text{Closed form, conjugacy}}}_{\text{convolution form with FFT}}\, d\boldsymbol{\omega_n}}_{L \text{ integral with quadrature method}}$$

Overall complexity : $\Longrightarrow$ $JKLT\log(T)$

Marginal distributions

Absence of target

$$p(u_n = 0 | \boldsymbol{Y_n}) = \prod_{l=1}^{L} p(u_n = 0)\Gamma(\bar{y}_{n,l} + \alpha_l^b)(T + \beta_l^b)^{-(\bar{y}_{n,l}+\alpha_l^b)}\gamma_l^{-1}$$

$$\gamma_l = \frac{\Gamma(\alpha_l^b)}{(\beta_l^b)^{\alpha_l^b}} \prod_{t=1}^{T} y_{n,l,t}$$

$$\bar{y}_{n,l} = \sum_{t=1}^{T} y_{n,l,t}$$

Marginal distributions

Presence of class k

$$p(u_n = k|\boldsymbol{Y_n}) = \sum_{d_n=1}^{T}\Big[\prod_{l=1}^{L} p(u_n = k)p(d_n)\mathcal{D}_{n,l,k}\gamma_l^{-1}\Big]\Big[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n)d\omega_{n,l}\Big]$$

Depth estimation

$$p(d_n|\boldsymbol{Y_n}) = \sum_{k=1}^{K}\Big[\prod_{l=1}^{L} p(u_n = k)p(d_n)\mathcal{D}_{n,l,k}\gamma_l^{-1}\Big]\Big[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n)d\omega_{n,l}\Big]$$

$$F_{n,l,k}(\omega_{n,l}, d_n) = \frac{\exp\{\sum_{t=1}^{T} y_{n,l,t}\ln[\omega_{n,l}\, Tg_l(t-d_n)+1]\}}{\omega_{n,l}^{1-\alpha_{k,l}^r}\{\beta_l^b + [T\,(1+\omega_{n,l}(1+\beta_{k,l}^r))]\}^{\alpha_l^b+\alpha_{k,l}^r+\bar{y}_{n,l}}}$$

Marginal distributions (Numerical challenges)

Presence of class k

$$p(u_n = k|\boldsymbol{Y_n}) = \sum_{d_n=1}^{T}\left[\prod_{l=1}^{L} p(u_n = k)p(d_n)\mathcal{D}_{n,l,k}\gamma_l^{-1}\right]\left[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n)d\omega_{n,l}\right]$$

Depth estimation

$$p(d_n|\boldsymbol{Y_n}) = \sum_{k=1}^{K}\left[\prod_{l=1}^{L} p(u_n = k)p(d_n)\mathcal{D}_{n,l,k}\gamma_l^{-1}\right]\left[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n)d\omega_{n,l}\right]$$

$$\mathrm{F}_{n,l,k}(\omega_{n,l}, d_n) = \frac{\exp\{\sum_{t=1}^{T} y_{n,l,t} \ln[\omega_{n,l}\, Tg_l(t-d_n)+1]\}}{\omega_{n,l}^{1-\alpha_{k,l}^r}\{\beta_l^b+[T\,(1+\omega_{n,l}(1+\beta_{k,l}^r))]\}^{\alpha_l^b+\alpha_{k,l}^r+\bar{y}_{n,l}}}$$

Solution

- Decouple the interlinked integrals by assuming different depth for each wavelength.

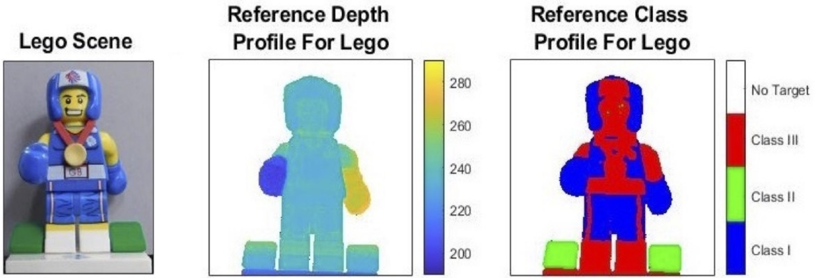
$$p(u_n = k|\boldsymbol{Y_n}) = \sum_{d_n=1}^{T}\left[\prod_{l=1}^{L} p(u_n = k)p(d_n)\mathcal{D}_{n,l,k}\gamma_l^{-1}\right]\left[\int_0^{\infty}\prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n)d\omega_{n,l}\right]$$
$$\approx \prod_{l=1}^{L}\sum_{d_{n,l}=1}^{T} p(u_n = k)p(d_{n,l})\mathcal{D}_{n,l,k}\gamma_l^{-1}\int_0^{\infty} F_{n,l,k}(\omega_{n,l}, d_{n,l})d\omega_{n,l}$$

- For depth estimation, consider point estimate instead of marginalising the SBR parameter, i.e., $p(d_n|\boldsymbol{Y_n}, \boldsymbol{\omega}_n^{\mathrm{map}})$ instead of $p(d_n|\boldsymbol{Y_n})$

Plan

Multispectral classification on real data

- 200 × 200 pixels
- Acquisition times per pixel and per wavelength : 40 ms
- Distance : 1.8 m
- Average Signal-to-Background Ratio (SBR) : $\omega = 66$ and $\omega = 1.3$
- Number of wavelengths and classes : L=4 and K=3



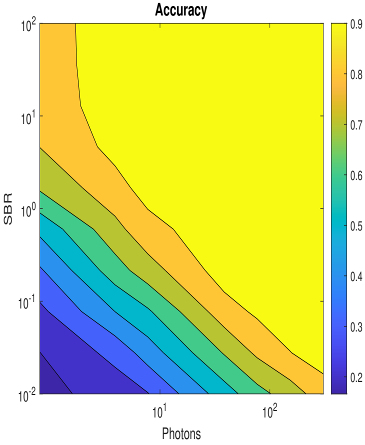
Evaluation criteria

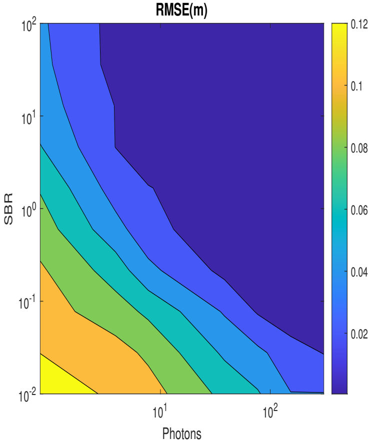
Criteria

- ***Depth*** : the root mean square error (RMSE) defined by $\mathrm{RMSE} = \sqrt{\frac{1}{N}||\boldsymbol{d}^{\mathrm{ref}} - \widehat{\boldsymbol{d}}||^2}$
- ***Class*** : Accuracy measure defined by $(\mathrm{ACC} = \frac{TP+TN}{TP+TN+FP+FN})$

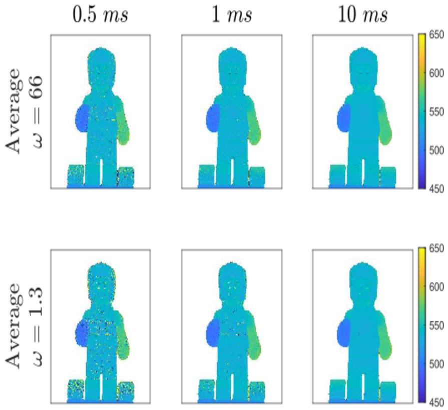
NB : The RMSE and accuracy evaluation will be done on the subsampled lego data (40×40 pixels).

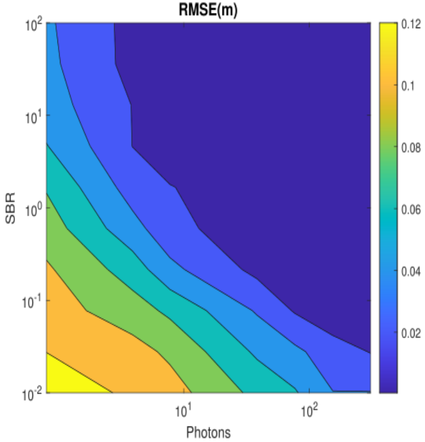
Results on real data : depth RMSEs and classification accuracy (40×40 pixels)



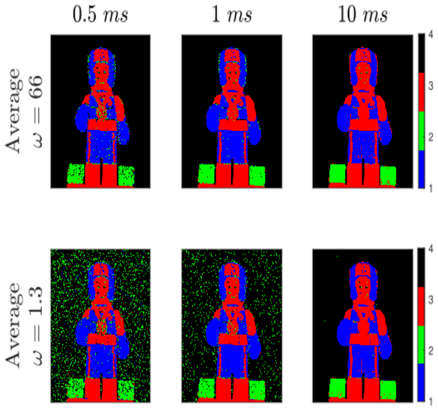


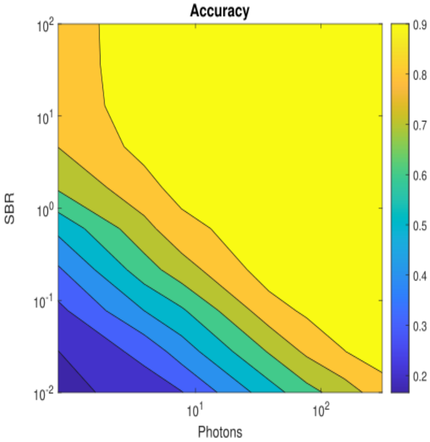
Results on real data : estimated depth and RMSE maps



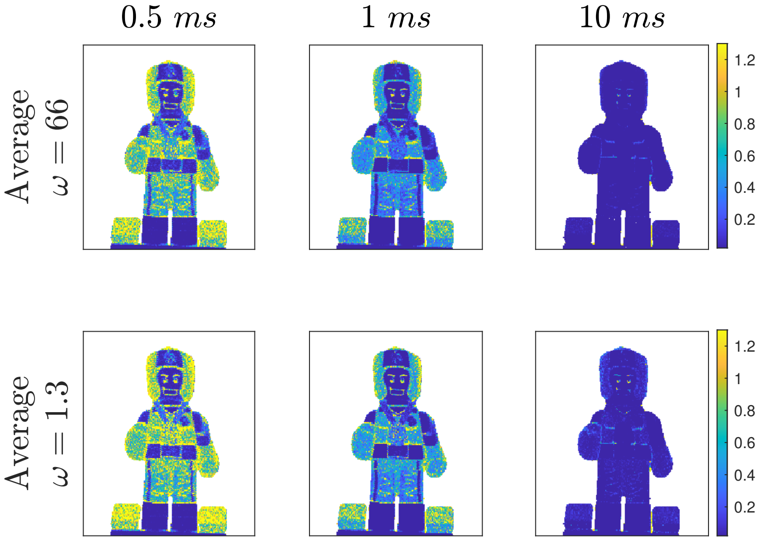


Results on real data : estimated classes lables and accuracy





Results on real data : the negative-log cumulative depth density



Plan

Conclusions

- A new multispectral classification model based on statistical framework.
- Fast and modular per-pixel approach that scale linearly with the number of pixels/wavelenghts.
- Possibility to perform a signature based target detection
- Possibility to get depth estimate and their uncertainty measure.

Future work

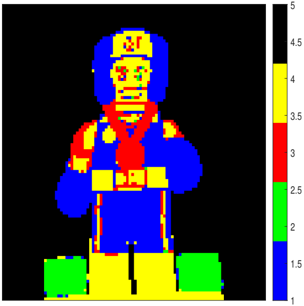
- Improve the data acquisition using an adaptive sampling approach
- Generalize to other challenging conditions (e.g. : multiple returns per-pixel, and non-uniform background)

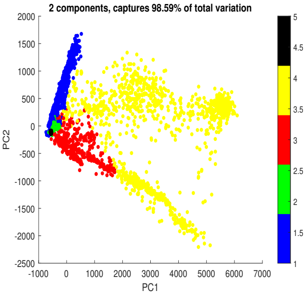
End

Thank you for your attention

Multispectral classification

2-PCA





Appendix II

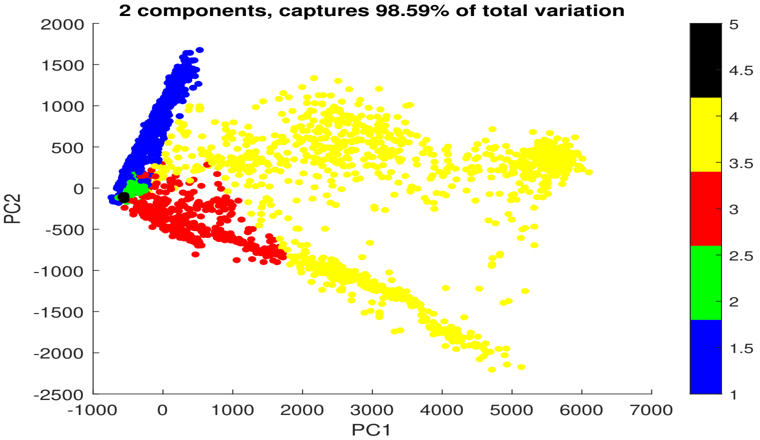


FIGURE – Projection of the pixels onto the 2 dominant principal components (2D).

Appendix III

Presence of class k

$$p(u_n = k|\boldsymbol{Y_n}) = \sum_{d_n=1}^{T} \Big[\prod_{l=1}^{L} p(u_n = k) p(d_n) \mathcal{D}_{n,l,k} \gamma_l^{-1} \Big] \Big[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n) d\omega_{n,l} \Big]$$

Depth estimation

$$p(d_n|\boldsymbol{Y_n}) = \sum_{k=1}^{K} \Big[\prod_{l=1}^{L} p(u_n = k) p(d_n) \mathcal{D}_{n,l,k} \gamma_l^{-1} \Big] \Big[\int_0^\infty \prod_{l=1}^{L} F_{n,l,k}(\omega_{n,l}, d_n) d\omega_{n,l} \Big]$$

$$\gamma_l = \frac{\Gamma(\alpha_l^b)}{(\beta_l^b)^{\alpha_l^b}} \prod_{t=1}^{T} y_{n,l,t} \,; \quad \mathcal{D}_{n,l,k} = \frac{\Gamma(\bar{y}_{n,l} + \alpha_l^b + \alpha_{k,l}^r)(T\beta_{k,l}^r)^{\alpha_{k,l}^r}}{\Gamma(\alpha_{k,l}^r)}$$

$$F_{n,l,k}(\omega_{n,l}, d_n) = \frac{\exp\{\sum_{t=1}^{T} y_{n,l,t} \ln[\omega_{n,l}\, T g_l(t - d_n) + 1]\}}{\omega_{n,l}^{1-\alpha_{k,l}^r} \{\beta_l^b + [T\,(1 + \omega_{n,l}(1 + \beta_{k,l}^r))]\}^{\alpha_l^b + \alpha_{k,l}^r + \bar{y}_{n,l}}}$$