Massachusetts Institute of Technology

Invariant Representation of Images under Change in Illumination

Sarah Luca[1], Fabio Anselmi[2], Georgios Evangelopoulos[2], Tomaso Poggio[2]
1: University of North Carolina at Asheville; 2: Center for Brains Minds and Machines and Massachusetts Institute of Technology.

Center for Brains, Minds & Machines

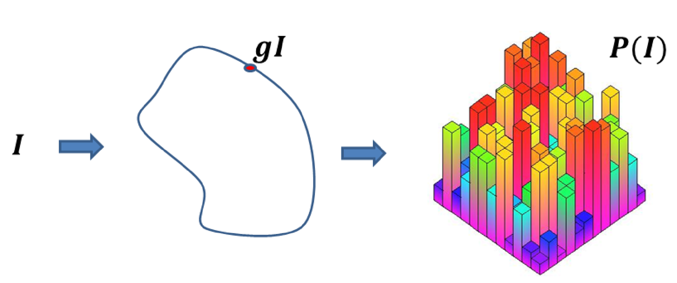
Introduction

∗ It has been shown [1] that an image representation based on a neurobiological model of simple-complex cells (Hubel and Wiesel model) is selective and invariant to group transformations and reduces the sample complexity of a learning task.
∗ This project aims to construct an invariant image representation under changes in illumination.
∗ We tested for illumination on two data sets: the SUFR data set [3] containing synthetic faces under 7 illumination conditions and the extended Yale Face Database B [4] containing images of human subjects under 64 illumination conditions.

Theoretical framework

G: finite compact group. *Equivalence relation* between images: $I \sim I' \Leftrightarrow \exists g_i \in G,\ I = g_i I'$. For an image I we have the image orbit, $O_I = \{g_1 I, \cdots, g_M I\}$, which is uniquely associated to a probability distribution P_I:

Theorem 1 $I \sim I' \Leftrightarrow O_I \sim O_{I'} \Leftrightarrow P_I = P_{I'}$



Implementing invariance and discriminability

∗ **Cramer-World Theorem**: probability distributions are uniquely determined by all of their one dimensional projections

$$P_I \equiv P_{I'} \Leftrightarrow P_{\langle I,t \rangle} \equiv P_{\langle I',t \rangle}$$

∗ **Group average**: Let G a finite compact group. The *group average* of any function $f : R \to R$ is

$$\bar{f}(x) \equiv \sum_i f(g_i x)$$

and we have $\bar{f}(x) = \bar{f}(g_i x),\ \forall g_i \in G$.

An invariant and selective signature

The following approximates the distribution of the values $\langle g_i I, t^k \rangle$ for one template t^k:

$$\mu_n^k(I) = \sum_i \left(\langle g_i I, t^k \rangle \right)^n$$

$\forall I, I'$ images we have:

Theorem 2 *The signature* $\mu(I) = (\mu_1^1(I), \cdots, \mu_N^K(I))$

- *is* **invariant** *i.e.* $\mu(g_i I) = \mu(I),\ \forall g_i \in G$
- *is* **selective (among classes)** *i.e.* $\mu(I) = \mu(I')\ \ iff\ \ I \sim I'$

If G is unitary, we have $\langle g_i I, t^k \rangle = \langle I, g_i^{-1} t^k \rangle \Rightarrow$ we have only one orbit of an arbitrary template to implement invariance of an image seen only once.

Framework for illumination transformations

∗ **Contrast Functions**: continuous, monotonic, and positive functions acting on an image.

We assume the contrast function $e : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^+$ has the following properties:

1. $e(x, 0) = 1$
2. $e(x, \zeta) e(x, \zeta') = e(x, \zeta \circ \zeta'),\quad \zeta, \zeta' \in \mathbb{R}_+$

where $\circ$ is a group composition. The transformation of an image I is:

$$I_\zeta(x) \equiv I(x, \zeta) = e(x, \zeta) I(x, 0).$$

and similarly for template t. We have the following signature for an image under change in illumination:

$$\mu_n^k(I) = \int_0^\infty \left(\langle I_0, t_\zeta^k \rangle \right)^n d\zeta$$

Illumination for the case of face images

For each of the datasets, we let I_0 be a face under one illumination condition and I_ζ be all other illumination conditions of the face. Our set of templates $\{t_\zeta^k\}$, contained faces under all illumination conditions with $t^k \neq I$.

Illumination for the case of Mondrian images

Let $e(x, m) = e^{mx}$ and $m = [-0.01, 0.01]$. For the illumination transformation,

$$e(x, m) I(x) = e^{mx} I(x) \equiv I_m(x),$$

and by theorems 1 and 2 we have the following invariant signature:

$$\mu_n^k(I) = \int_0^\infty \left(\langle I_0, t_m^k \rangle \right)^n dm$$

Methods

∗ **Invariance**: We calculated the euclidean distance between images and their transformations. For invariance we tested:

$$||\mu(I_0) - \mu(I_\zeta)||^2 \sim 0$$

where we fixed 1 template, with $\mu(I) \in \mathbb{R}^{n \times k}$, and averaged the euclidean distance over images and their transformations.

∗ **Selectivity**: We calculated the euclidean difference between different images. For selectivity we tested:

$$||\mu(I_0^i) - \mu(I_0^j)||^2 \nsim 0,\ \ i \neq j$$

where we averaged over many templates.

∗ **Control**: We defined a "fake orbit" as a set of randomly selected images from the datasets.

Results

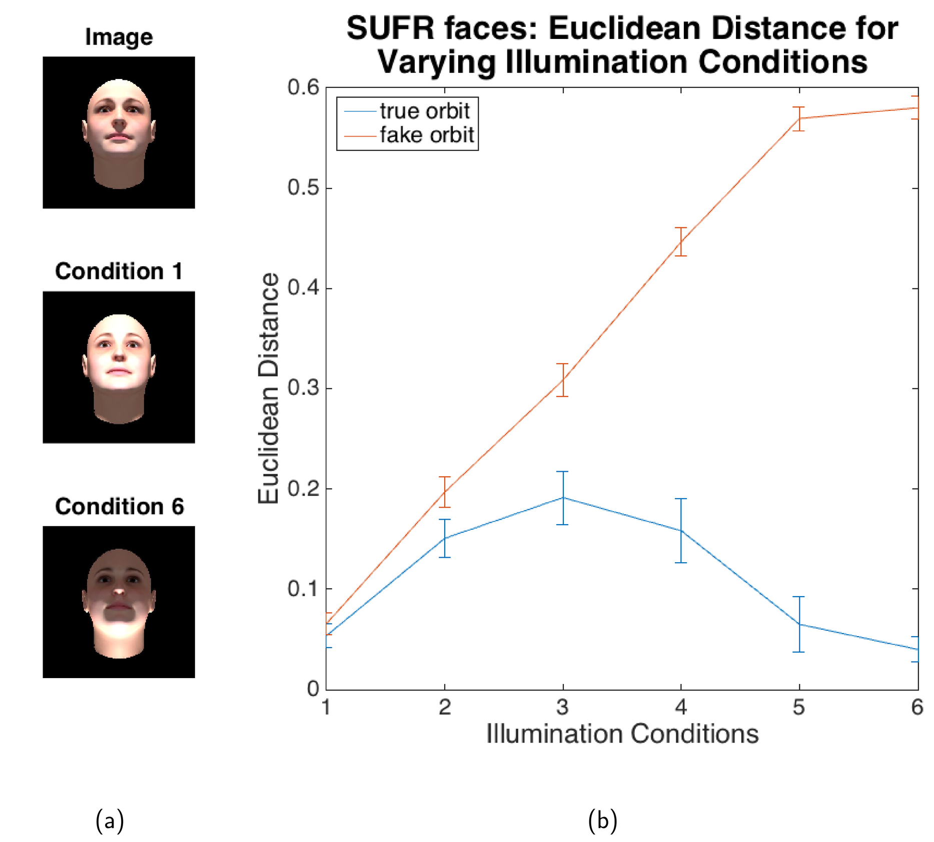


Figure 1: (a) Example of a face from SUFR dataset and two illumination conditions. (b) Euclidean distances between faces within the same orbit for the true and fake orbits.

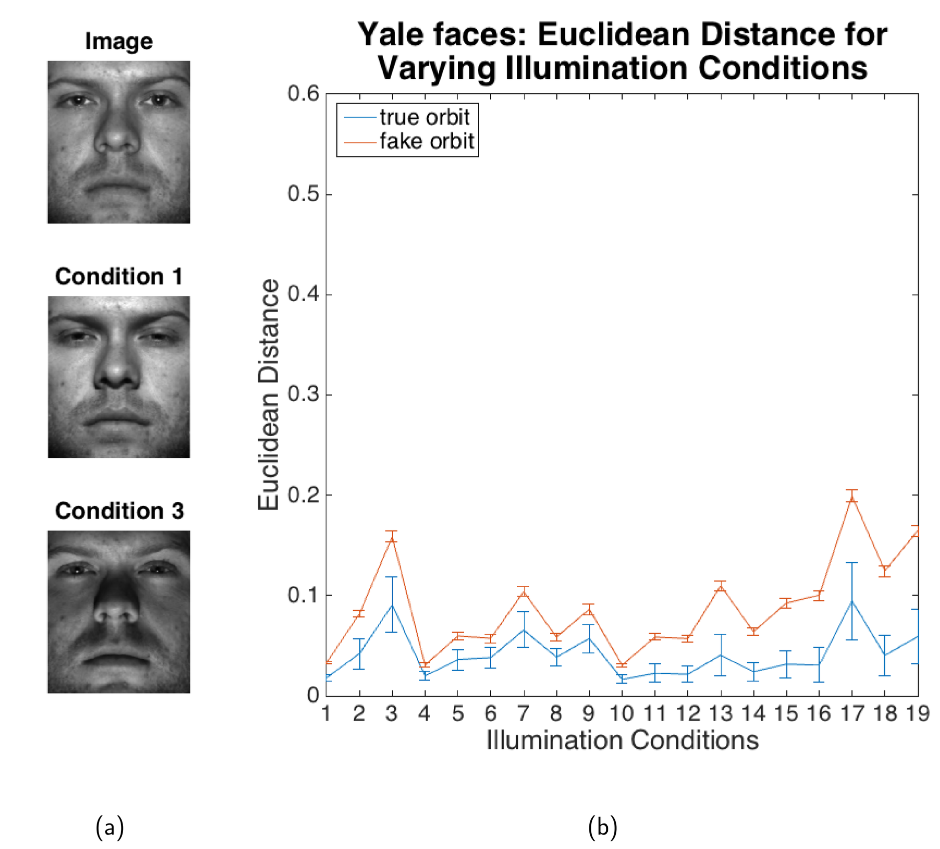


Figure 2: (a) Example of a face from YaleFace dataset and two illumination conditions. (b) Euclidean distances between faces within the same orbit for the true and fake orbits.

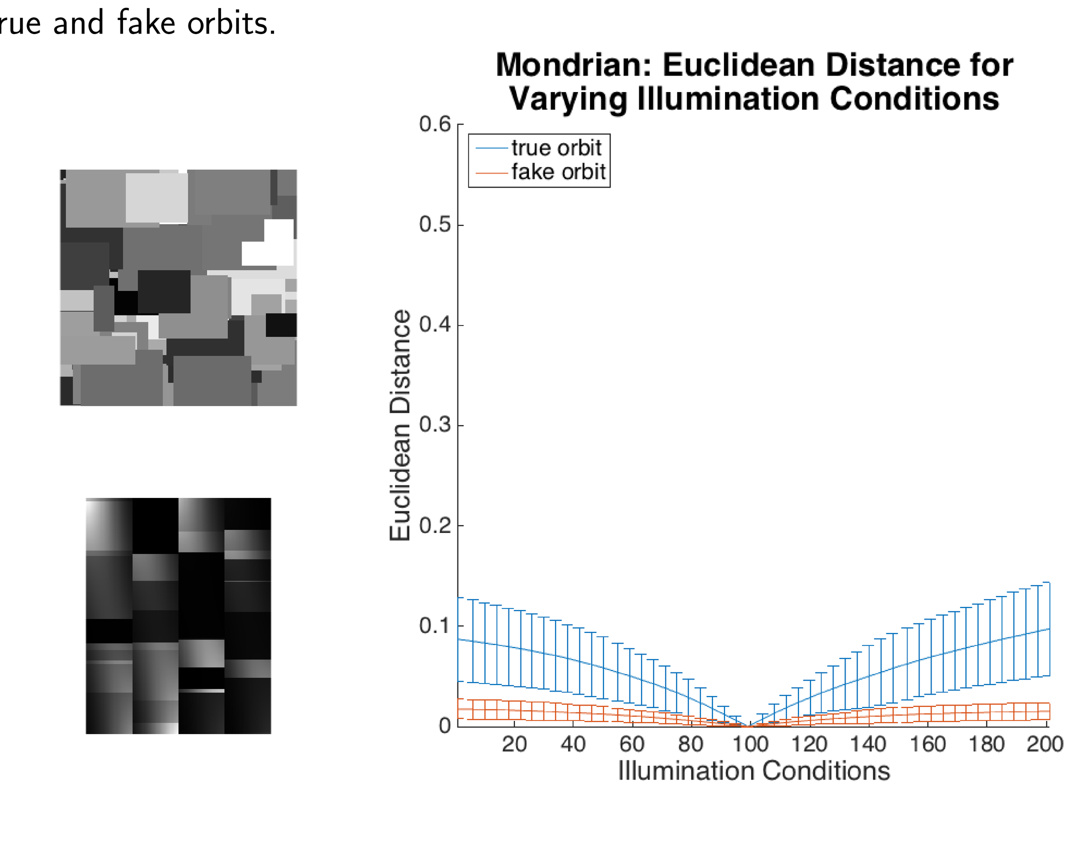


Figure 3: (a) Example of a Mondrian image (top) and strips of the Mondrian under different illumination conditions (bottom). (b) Euclidean distance between images within the same orbit for the true and fake orbits.

Discussion

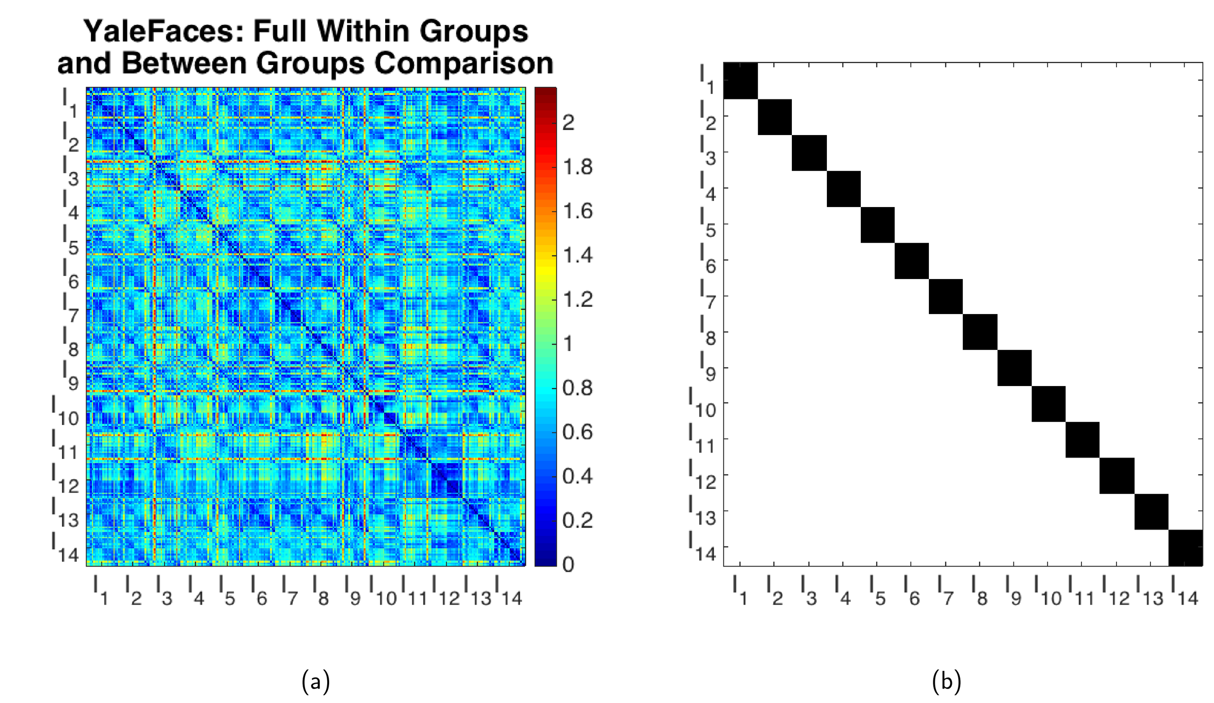


Figure 4: (a) Within groups and between groups comparison of faces. Within groups measured the difference between faces and each of their illuminations (I and each $g_i I$). Between groups measured the difference between different faces (I_i and I_j). (b) The ground truth distance matrix. Each block is the within groups comparison, and white is the between groups comparison.

∗ **Invariance**: Results suggest improved invariance compared to the control for face images but not for Mondrian images.
∗ **Selectivity**: Results were inconclusive (see Figure 4).

∗ **Possible explanations**:

1. normalization of signatures/other distance metrics
2. limited orbit size (number of transformed images)
3. assumptions made about illumination transformations/models

Summary

We considered a framework for invariance under illumination transformations of images and tested the theory on real and artificial data. For face images with a limited number of orbit samples, we observed invariance but not selectivity. Future work will include finer illumination models and experiments with more controlled, synthetic light conditions.

Acknowledgements

This material is based upon work supported by the Center for Brains, Minds and Machines (CBMM), funded by NSF STC award CCF-1231216.

References

1. Anselmi F., Leibo J.Z., Rosasco L., Mutch J., Tacchetti A., Poggio T., Unsupervised Learning of Invariant Representations, arXiv:1311.4158.
2. Hurlbert A.C. and Poggio T., Synthesizing a Color Algorithm from Examples (1988), Science 239.4839 (1988): 482-485.
3. Leibo J. Z., Liao Q., and Poggio T., Subtasks of Unconstrained Faces Recognition (2014). 9th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications. (VISAPP). Lisbon, Portugal.
4. Lee K.C., Ho J. and Kriegman D., Acquiring Linear Subspaces for Face Recognition under Variable Lighting (2005). IEEE Trans. Pattern Anal. Mach. Intelligence. 27(5) pp 684-698.