A Computational Model for the Biological Underpinnings of Infant Vision and Face Recognition

Laura A. Mueller and Amir Assadi
Mathematics Department, University of Wisconsin-Madison

Abstract

While vision science has made great strives in a multitude of directions, little is known about computational models of infant vision. Upon birth, brains contain all necessary neurons but lack many connections. Within a few months after birth, infants can easily recognize familiar faces and comprehend expressions, suggesting there exists a predisposition for the human face. In this project, an artificial neural network is created in order to model and understand how an infant's brain processes visual information and hardwires the necessary connections to efficiently recognize faces. The neural network model will learn, as an infant does to filter facial images in search of salient features that distinguish the face from other objects.

Summary

Regarding face recognition by infants, the challenge is to learn how an infant's brain learns to efficiently process visual information. An infant's range of focus is short for several months, and faces enter into this range more often than any other scene. Within a few months, infants can differentiate faces from other objects and are able to recognize a known face from that of a stranger. The main purpose of this project is to examine how an infant learns to encode visual stimuli efficiently through an artificial neural network simulation. The repetition of face stimulus in conjunction with the relatively short time it takes for recognition to occur suggests there may be some regularity among the stimuli. Rather than performing physically invasive experiments, a neural network will serve as a model for an infant's brain, learning to distinguish a fact from other objects given little or no previous knowledge of face structure.

Our theoretical basis for our model is based on information theory, considering the way in which a brain filters and represents a facial image in a mathematically comprehensive form. We hypothesize the biologically that the brain doesn't need to study details of each individual face in order to correctly categorize it from other objects. The repetition of face stimulus in conjunction with the relatively short time it takes for recognition to occur suggests there may be some regularity among the stimuli. The brain may have to thoroughly examine a few faces in order to perceive regularities, and then only look for those regularities in other faces. These regularities are salient facial features such as eyes, nose, and mouth. It's unknown what features primarily trigger the brain, so several features will be fed to the neural network in attempt to determine those features carrying as much information as an entire face. Through the training process, the neural network will become as complex as an infant's brain, securing neuron connections as visual stimuli is processed.

The network will be trained with frontal view face images, prompting it to rigorously examine the image and associate face with the image. A database of images, collected from a digital camera and magazine clippings, are manipulated via software to eliminate background variations and narrow focus to the face. The images are transformed from square matrices to vector form and arranged in one matrix, each column corresponding to a face vector. After normalizing the matrix and performing a transform such as discrete cosine or Fourier, the correlation matrix is computed by multiplying the matrix by its transpose. Principal component analysis (PCA) is performed on the correlation matrix to determine the eigenvalues and eigenvectors. From the list of eigenvalues, the percentage of information is computed, where information is the ratio of the first n eigenvalue to the sum of all eigenvalues. This value indicates how many eigenvectors carry a specified percentage of the entire set of images. To produce what some term an eigenface, the matrix of face vectors is multiplied by the principal component. Mathematically this results in projecting each face vector onto the principal component, the result of the multiplication being one vector representing the primary coordinate in a new space. This vector is reshaped into the original dimensions of the face image. The resulting image is what some term an eigenface, a general face image lacking clarity of specific facial features. If there are multiple principal components, multiple eigenfaces are needed to represent the determined percentage of information. The neural network will be trained on these eigenfaces, with an eigenface representing the result of the first step of filtering performed by the brain's visual cortex.

After the training session, the neural network will be tested with images from portions of the face. For example, one image may show the left half of the face or another display the nose and eyes. These face chunks will also undergo PCA before being presented to the neural network, whereby the network learns to recognize the images as portions of a face either by the partial image alone or by piecing it together with other partial images. Efficiency may be evaluated as the network requires fewer partial images to construct a face, meaning the features trigger recognition. The triggering suggests a hardwiring of neuron connections between the eye and visual cortex. In this project, the computational model presented to simulate the learning process of infant vision closely resembles the biological mechanisms of the brain that process complex visual stimuli by recognizing patterns and regularities.