Motion Contrast Classification Is a Linearly Nonseparable Problem

Alireza S. Mahani
amahani@wustl.edu
Ralf Wessel
rw@wuphys.wustl.edu
Physics Department, Washington University,
St. Louis, MO 63130, U.S.A.

Sensitivity to image motion contrast, that is, the relative motion between different parts of the visual field, is a common and computationally important property of many neurons in the visual pathways of vertebrates. Here we illustrate that, as a classification problem, motion contrast detection is linearly nonseparable. In order to do so, we prove a theorem stating a sufficient condition for linear nonseparability. We argue that nonlinear combinations of local measurements of velocity at different locations and times are needed in order to solve the motion contrast problem.

1 Introduction

Many neurons in the visual pathways of vertebrates, such as cells in the middle temporal area of primates (Allman, Miezin, & McGuinness, 1985) and cells in the deep layers of the avian optic tectum (Frost & Nakayama, 1983), are sensitive to image motion contrast. Motion contrast sensitivity is believed to play a major role in important perceptual functions such as figure-ground segregation (Nakayama, 1985). Here we illustrate that as a classification problem, motion contrast detection is linearly nonseparable.

Section 2 defines the concept of linear separability for a binary classification problem. Section 3 offers a sufficient condition for a problem to be linearly nonseparable. Section 4 uses the result of section 3 to prove that motion contrast detection is linearly nonseparable. Section 5 generalizes the nonseparability of the motion contrast problem to include a certain type of nonlinear preprocessing. In section 6 we discuss what happens when image intensity is used as input for calculating motion contrast. Section 7 offers concluding remarks.
2 Defining Linear Separability

Suppose that we have a set of $d$-dimensional samples $\mathbf{x}_1, \ldots, \mathbf{x}_n$, $n_1$ of which are in the subset $\Omega_1$ and $n_2 = n - n_1$ in the subset $\Omega_2$. If we form a linear combination of the components of $\mathbf{x}_i$, we obtain the scalar

$$y_i = \mathbf{w}^T \mathbf{x}_i.$$  \hspace{1cm} (2.1)

This operation produces a set of $n$ samples $y_1, \ldots, y_n$ from the input samples $\mathbf{x}_1, \ldots, \mathbf{x}_n$. The two subsets $\Omega_1$ and $\Omega_2$ are linearly separable if there exists a vector $\mathbf{w}$ and a scalar $\gamma$ such that

$$\forall \mathbf{x}_i \in \Omega_1, \ y_i > \gamma \quad \text{and} \quad \forall \mathbf{x}_i \in \Omega_2, \ y_i < \gamma,$$  \hspace{1cm} (2.2)

with $y_i$’s defined in equation 2.1. In other words, two classes are linearly separable if they can be completely separated by a hyperplane in the feature space. If such a hyperplane cannot be found, the two classes are called linearly nonseparable. For more on linear separability, see Duda, Hart, and Stork (2001).

3 A Sufficient Condition for Linear Nonseparability

Given the above definition of linear separability, it is easy to prove the following theorem.

**Theorem 1.** Two classes are linearly nonseparable if we can find two samples from each class, $\bar{x}_1, \bar{x}_2 \in \Omega_1$ and $\bar{x}_3, \bar{x}_4 \in \Omega_2$, and two numbers $\alpha$ and $\beta$, such that

$$\alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2 = \beta \bar{x}_3 + (1 - \beta) \bar{x}_4,$$  \hspace{1cm} (3.1)

$$0 \leq \alpha, \beta \leq 1.$$  \hspace{1cm} (3.2)

**Proof.** If the two classes are separable, there are $\mathbf{w}$ and $\gamma$ such that $\mathbf{w}^T \bar{x}_1 > \gamma$, $\mathbf{w}^T \bar{x}_2 > \gamma$, $\mathbf{w}^T \bar{x}_3 < \gamma$, and $\mathbf{w}^T \bar{x}_4 < \gamma$. From these four inequalities, we conclude that $\mathbf{w}^T (\alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2) > (1 - \alpha + \alpha) \gamma = \gamma$ and $\mathbf{w}^T (\beta \bar{x}_3 + (1 - \beta) \bar{x}_4) < (1 - \beta + \beta) \gamma = \gamma$. These in turn indicate that $\alpha \bar{x}_1 + (1 - \alpha) \bar{x}_2 \neq \beta \bar{x}_3 + (1 - \beta) \bar{x}_4$, which leads to contradiction.

As a special case, we can have $\alpha = \beta = 0.5$. Equation 3.1 will now read $\bar{x}_1 + \bar{x}_2 = \bar{x}_3 + \bar{x}_4$; the sums of the two samples in each class are the same.
Motion contrast No motion contrast

<table>
<thead>
<tr>
<th></th>
<th>Motion contrast</th>
<th></th>
<th>No motion contrast</th>
</tr>
</thead>
<tbody>
<tr>
<td>$x_1$</td>
<td><img src="image1.png" alt="image" /></td>
<td></td>
<td><img src="image2.png" alt="image" /></td>
</tr>
<tr>
<td>$x_2$</td>
<td><img src="image3.png" alt="image" /></td>
<td></td>
<td><img src="image4.png" alt="image" /></td>
</tr>
<tr>
<td>$x_1 + x_2$</td>
<td><img src="image5.png" alt="image" /></td>
<td></td>
<td><img src="image6.png" alt="image" /></td>
</tr>
</tbody>
</table>

Figure 1: Pairs of points in feature space for motion contrast versus no motion contrast (top two rows). Arrows indicate the velocity of each dot, which is either $v$ or $-v$. Both pairs have the same sum. This, according to theorem 1, is sufficient to prove the nonseparability of this classification problem.

4 Linear Nonseparability of the Motion Contrast Classification Problem

Using theorem 1, we can easily demonstrate the nonseparability of motion contrast detection problem for two dots. Figure 1 illustrates a case where the sums of two samples from the two classes (with and without motion contrast) are equal. In terms of equation 3.1, we have $\tilde{x}_1 = (v, -v), \tilde{x}_2 = (-v, v)$ and $\tilde{x}_3 = (v, v), \tilde{x}_4 = (-v, -v)$, and therefore $\tilde{x}_1 + \tilde{x}_2 = \tilde{x}_3 + \tilde{x}_4 = (0, 0)$. According to theorem 3.1, this proves that the problem is linearly nonseparable. The same argument can be applied to an object against a background, ignoring the occlusion and edge effects.

5 Point-Wise Static Preprocessing

We have observed that a simple linear transformation of the input cannot separate the coherent and noncoherent cases. It can easily be shown that even if this linear transformation is preceded by a point-wise static non-linear transformation, separation of the two classes of inputs still cannot be achieved. A point-wise static transformation, if considered in isolation, is nothing more than a scalar function of a scalar variable: $g(v)$. The qualifiers “static” and “point-wise” are meant to emphasize that as a transformation applied to a dynamic scalar field, such as a time-varying image intensity field, it does not integrate (in the general sense) the values of the function over time or space to produce the output.

Again, consider the two-dot scenario with $\tilde{x}_1 = (v, -v), \tilde{x}_2 = (-v, v)$ (motion contrast) and $\tilde{x}_3 = (v, v), \tilde{x}_4 = (-v, -v)$ (no motion contrast). If we apply $g( )$ to each of these samples, we find

$$\tilde{x}_1 = (g(v), g(-v)), \quad \tilde{x}_2 = (g(-v), g(v)),$$

$$\tilde{x}_3 = (g(v), g(v)), \quad \tilde{x}_4 = (g(-v), g(-v)).$$

(5.1)
Obviously, we still have $\vec{x}_1' + \vec{x}_2' = \vec{x}_3' + \vec{x}_4'$. Therefore, the problem remains linearly nonseparable.

### 6 Image Intensity Field as Input

So far we have assumed that motion contrast detection uses the velocity field as input. It is indeed possible to detect motion contrast by directly using the image intensity. For example, consider two dots moving at velocities $\vec{v}_1$ and $\vec{v}_2$. A reasonable definition of the motion contrast (MC) of this arrangement is $MC = |\vec{v}_1 - \vec{v}_2|$. Clearly, $\vec{v}_1 = d\vec{r}_1/dt$ and $\vec{v}_2 = d\vec{r}_2/dt$ with $\vec{r}_1$ and $\vec{r}_2$ being the displacement vectors for the two objects. We can now rephrase motion contrast as $MC = |\vec{v}_1 - \vec{v}_2| = |d\vec{r}_1/dt - d\vec{r}_2/dt| = |d(\vec{r}_1 - \vec{r}_2)/dt|$. The last expression shows that we can detect motion contrast between two moving dots by monitoring the change in their distance rather than calculating their velocities first. Dellen, Clark, and Wessel (2004) have suggested a more general method for detecting motion contrast that does not need velocity information. Is motion contrast detection still linearly nonseparable if we use intensity instead of velocity information? The answer is yes. An argument very similar to that presented here can be made for the intensity-based approach. A point-wise static nonlinearity applied to the intensity field does not change the linear nonseparability either. This is because if we consider two widely separated objects in motion and apply the point-wise static transformation to the resulting intensity field, we arrive at two objects with possibly distorted profiles moving at the same velocities as the original intensity field, and the distortion is independent of velocities. Therefore, if the original problem is linearly nonseparable, so is the new one.

### 7 Discussion

In summary, we showed that motion contrast detection cannot happen through a linear integration of local inputs, even if these local inputs have been preprocessed by a static point-wise transformation (linear or non-linear).

In the primate visual cortex, the earliest area with motion contrast sensitive neurons is the middle temporal cortex (area MT). The main pathway leading from the retina to area MT passes through the lateral geniculate nucleus (LGN) and primary visual cortex (area $V_1$). The presence of multiple stages between where the image intensity map is represented and where motion contrast is represented has naturally led many researchers to assume that in primates, velocity field is calculated prior to the computation of motion contrast. As a result, in building cortical models of image motion processing, most of the attention and effort has been devoted to estimating local image velocity (Poggio & Reichardt, 1973; Adelson & Bergen, 1985;
Simoncelli & Heeger, 1998). All these models require nonlinear processing of image intensity in order to estimate local image velocity. The highly nonlinear (phase-invariant) response of complex cells in area V₁ (Skottun et al., 1991) is an important example of such nonlinearities. In this article, however, we began by assuming that the velocity field was provided as input to the motion contrast detection problem. Therefore, our conclusion that nonlinearities other than the point-wise static type are needed to perform the motion contrast detection task is independent of the well-known fact that estimation of velocity field requires nonlinear processing.

Simoncelli and Heeger (1998) use a nonlinear operation called divisive normalization to reduce the dependency of the response of their model MT neurons on image contrast. In its original form, this nonlinearity is restricted to the classical (center) receptive field of the neuron. The authors have suggested, however, that a similar normalization that pools neurons with opposite-direction preferences in the center and surround areas may result in motion contrast sensitivity. While our result cannot confirm the above hypothesis regarding the origin of motion contrast sensitivity in the MT neurons, they are certainly consistent with each other, as divisive normalization involves nonlinear interactions across space and therefore is not point-wise. Physiologically plausible ways of implementing such nonlinearities, using inhibitory feedback, have been suggested and (indirectly) tested (for example, see Carandini, Heeger, & Movshon, 1997).

In the nonmammalian visual system, deep tectal neurons with motion contrast sensitivity are suspected to be only one synapse away from the retina (Luksch, 2003). While this observation does not exclude the possibility of other inputs to such cells carrying the local velocity information, it certainly motivates research on possible ways to extract motion contrast information directly from the image intensity (Dellen et al., 2004). In this article, we argued that such a direct solution would still require nonlinear processing of a similar nature to that required while using local velocity as input. Anatomical (Luksch, 2003) and electrophysiological (Wang, 2003) studies suggest the existence of a number of other neural structures and pathways, both within and outside the optic tectum, with a potentially nonlinear role in shaping the complex response properties of deep tectal neurons, including their motion contrast sensitivity. Examples are the network of horizontal neurons in the optic tectum and the feedback loop between the optic tectum and its midbrain satellite, nucleus isthmi.

Acknowledgments

We thank Anders E. Carlsson and the anonymous reviewers for their critical reading of the manuscript. This work was supported by grants from the Whitehall Foundation and the McDonnell Center for Higher Brain Function to R.W.
References


Received April 23, 2004; accepted January 27, 2005.