Appendix

Derivations for the following formulas are found in Munoz et al [1].

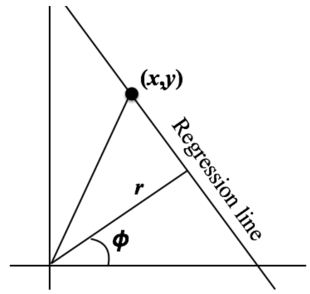

Figure 1: *Regression Line*

Munoz et al use for their derivations the normal form of the equation of a straight line, in which the two parameters of the fitted line, φ and r, are the polar coordinates of the intersection of the fitted line and another line from the origin normal to it, as in Figure 1. For a point (x,y) on the regression line, the usual slope-intercept form of the equation of the orthogonal regression line, $y = mx + b$, is given by $r = x \cos\varphi + y \sin\varphi$, which rearranges to

$$y = \frac{-1}{\tan\varphi} x + \frac{r}{\sin\varphi}$$

The ellipse axis angle is calculated as follows. In order to avoid denominators equal to zero, if the denominator is less than .001 then it is replaced by .001. If $\varphi = 0$ then it is replaced by 0.01 radians.

$$\varphi(t) = 0.5\,\mathrm{Arctan}[\frac{-2\sum_{i=1}^{d} w_i(t)(x_i(t) - \bar{x}(t)(y_i(t) - \bar{y}(t))}{\sum_{i=1}^{d} w_i(t)(y_i(t) - \bar{y}(t))^2 - (x_i(t) - \bar{x}(t))^2)}]$$

For $j = 1,2$, the slopes m_j and intercepts b_j for the two axes lines $y_j(t) = m_j(t)\, x_j(t) + b_j(t)$ are calculated. One of these lines minimizes fitting errors and becomes the major ellipse axis and the other maximizes the errors and is the minor axis:

$$m_1(t) = \frac{-1}{\tan\varphi(t)}, \quad b_1(t) = \bar{y}_1 + \frac{\bar{x}_1}{\tan\varphi(t)}$$

$$m_2(t) = \tan\varphi(t), \quad b_2(t) = \bar{y}_2(t) - \bar{x}_2(t) \cdot \tan\varphi(t)$$

As in Munoz et al, the standard deviations are calculated as follows:

$$SD_1(t) = \sqrt{\frac{\sum_{i=1}^{d} [w_i(t)(x_i(t) - \bar{x}_i(t))\cos\varphi(t) + w_i(t)(y_i(t) - \bar{y}_i(t))\sin\varphi(t)]^2}{\sum_{i=1}^{d} w_i(t)}}$$

$$SD_2(t)=\sqrt{\frac{\sum_{i=1}^{d}[w_i(t)(x_i(t)-\bar{x}_i(t))\sin\varphi(t)-w_i(t)(y_i(t)-\bar{y}_i(t))\cos\varphi(t)]^2}{\sum_{i=1}^{d}w_i(t)}}$$

For each time t the semi-major axis is calculated as the standard deviation, either SD_1 or SD_2, of the orthogonal regression line which has the larger value, and the semi-minor axis is the standard deviation of the line perpendicular to the first line, which has the smaller standard deviation value. Thus, the larger and smaller of the two standard deviations above are used, with a desired factor such as 1.5, as semi-major(t) and semi-minor(t) axes respectively, in the animation of the ellipse associated with the dancers' movements.

In order that the ellipse not disappear too often in the animation, for example, when one dancer is on stage or the dancers are clustered in one location, if the major axis < 0.5 then it is set to 0.5 and if the minor axis < 0.25 then it is set to 0.25. The ellipse at time t is calculated using Mathematica's ellipsoid function and interpolation function. The interpolation functions we have used for the axes and other calculations have been linear, as higher order functions produced artifacts unlike the actual dancers' movements. Mathematica's resolution within its animation command smooths out the sharp turns that would otherwise appear due to the use of linear interpolation. Calculations are as follows, where the linear factor has been set to 1.5, which judgmentally best encompasses dancer positions, and the exponential factor to 1 for our initial calculations.

Center: interpolated Center of Attention functions for x and for y coordinates at time t.
Major axis: (linear factor)(interpolated major axis function at time t)$^{\text{exponential factor}}$
Minor axis: (linear factor)(interpolated minor axis function at time t)$^{\text{exponential factor}}$

Dancers: The disks for dancers are created using a varying radius dependent on the interpolated weight function at time t and using colors designed to contrast with each other. The disk radius r_i for the ith dancer is calculated using the formula $r_i = 0.15 + 0.30\frac{w_i(t)}{\sum_{i=1}^{d} w_i(t)}$ so the radii will be between 0.15 and 0.45. The color of the ith disk is given by $H_i(h,s,b)$, where h is hue, s is saturation, and b is brightness, and $H_i=(\frac{i}{d},0.5,\frac{i}{d})$. We also calculate the instantaneous average weighted distance of the dancers from the CA along with the average instantaneous standard deviation of this distance. If $D((a_i,b_i),(a_j,b_j))$ is the Euclidean distance between points (a_i,b_i) and (a_j,b_j) then the weighted average distance at time t, $D_{avg}(t)$ is calculated as follows and incorporated into the calculation of the weighted standard deviation:

$$D_{Avg}(t)=\frac{\sum_{i=1}^{d}w_i(t)\cdot D((x_i(t),y_i(t)),(\bar{x}(t),\bar{y}(t)))}{\sum_{i=1}^{d}w_i(t)},$$

$$\text{Weighted standard deviation at time } t = \frac{\sum_{i=1}^{d}w_i(t)\cdot D((x_i(t),y_i(t)),(\bar{x}(t),\bar{y}(t))-D_{Avg}(t))^2}{\sum_{i=1}^{d}w_i(t)}$$