Expected Confusion as a Method of Evaluating Recognition Techniques

Georgia Institute of Technology Technical Report: GIT.GVU-01-10

Aaron F. Bobick

GVU Center/College of Computing
Georgia Tech
Atlanta, GA 30332
afb@cc.gatech.edu

Amos Y. Johnson

Electrical and Computer Engineering
Georgia Tech
Atlanta, GA 30332
amos@cc.gatech.edu

Abstract

We derive an expected confusion metric, as opposed to reporting percent correct with a limited database, as a method to evaluate recognition techniques. This metric allows us to predict how well a given feature vector will filter identity in a large population. Our expected confusion is the ratio of the average individual variation of a feature vector to that of the population variation of the feature vector. We evaluate our gait-recognition technique [2] that recovers static body and stride parameters of walking subjects with the expected confusion metric to demonstrate its use.

1. Introduction

Perhaps the most significant limitation in recognition techniques is the manner in which results are reported. Even though the size of training/testing subject databases are typically much smaller then a population, results are reported as percent correct. That is, on how many trials could the system correctly recognize the individual by choosing its best match. Such a result gives little insight as to how the technique might scale when the database contains hundreds or thousands or more people.

As opposed to reporting percent correct, we will establish the uncertainty reduction that occurs when a measurement is taken. For a given measured property, we establish the density of the overall population. To do so requires only enough subjects such that the density approaches some stable estimate. Next we determine the average or expected variation in the measurement when applied to a given individual. Essentially the ratio of the volume of the two densities gives an indication of how "good" a biometric the measurement is.

The reminder of this paper is as follows: we describe the expected confusion metric, and demonstrate its use on our gait-recognition technique that recovers static body and stride parameters of walking subjects.

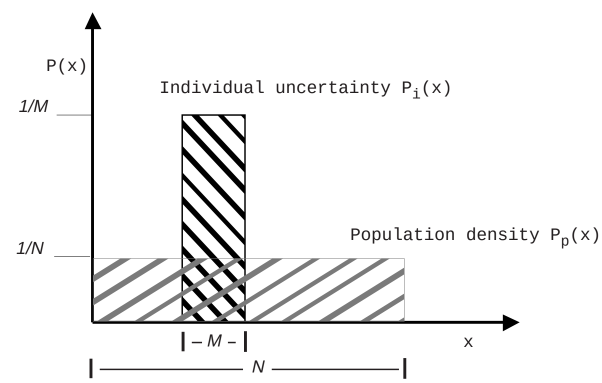


Figure 1: Uniform probability illustration of how the density of the overall population compares to the the individual uncertainty after the measurement is taken. In this case the remaining confusion — the percentage of the population that could have given rise to the measurement — is M/N.

2. Expected Confusion

As mentioned our goal is not to report a percent correct of identification. To do so requires us to have an extensive database of thousands of individuals being observed under a variety of conditions. Rather, our goal is to characterize a particular measurement as to how much it reduces the uncertainty of identity after the measurement is taken.

Many approaches are possible. Each entails first estimating the probability density of a given property vector $\mathbf{x}$ for an entire population $P_p(\mathbf{x})$. Next we must estimate the uncertainty of that property for a given individual once the measurement is known $P_I(\mathbf{x}|\eta = \mathbf{x}_0)$ (interpreted as what is the probability density of the true value of the property $\mathbf{x}$ after the measurement η is taken). Finally, we need to express the average reduction in uncertainty or the remaining confusion that results after having taken the measurement.

Information theory argues for a mutual information [1] measure:

$$I(\mathbf{X};\mathbf{Y}) = H(\mathbf{X}) - H(\mathbf{X}|\mathbf{Y}). \tag{1}$$

where $H(\mathbf{X})$ is the entropy of a random variable $\mathbf{X}$ defined by

$$H(\mathbf{X}) = -\int_x p(x)\ln p(x),$$

and $H(\mathbf{X}|\mathbf{Y})$ is the conditional entropy of a random variable $\mathbf{X}$ given another random variable $\mathbf{Y}$ defined by:

$$H(\mathbf{X}|\mathbf{Y}) = -\int_{x,y} p(x,y)\ln p(x|y).$$

For our case the random variable $\mathbf{X}$ is the underlying property (of identity) of an individual before a measurement is taken and is represented by the population density of the particular metric used for identification. The random variable $\mathbf{Y}$ is an actual measurement retrieved from an individual and is represented by a distribution of the individual variation of an identity measurement. Given these definitions, the uncertainty of the property (of identity) of the individual given a specific measurement, $H(\mathbf{X}|\mathbf{Y})$, is just the uncertainty of the measurement, $H(\mathbf{Y})$. Therefore the mutual information reduces to:

$$I(\mathbf{X};\mathbf{Y}) \equiv H(\mathbf{X}) - H(\mathbf{Y}). \tag{2}$$

Since the goal of gait recognition is filtering human identity this derivation of mutual information is representative of filtering identity. However, we believe that a better assessment (and comparable to mutual information) of a metric's ability to filter identity is the expected value of the percentage of the population eliminated after the measurement is taken. This is illustrated in Figure 1. Using a uniform density for illustration we let the density of the feature in the population P_p be $1/N$ in the interval $[0,N]$. The individual density P_i is much narrower, being uniform in $[x_0 - M/2, x_0 + M/2]$. The confusion that remains is the area of the density P_p that lies under P_i. In this case, that confusion ratio is M/N.

When the densities are not uniform, the situation is a little more complicated. We still need the area under P_p that falls under P_i, but that area needs to be weighted by the value of P_i. Consider a little strip Δx under both P_p and P_i. The incremental area in that strip (with respect to P_p) in the neighborhood of x is just $P_p(x)\cdot\Delta x$. Suppose we have taken a measurement of value η (a slight abuse of notation from above). If we assume that P_i is zero mean noise about the measurement, then the absolute scaled incremental area ΔA is:

$$\Delta A^*(\eta,x) = P_p(x)\,P_i(\eta - x)\,\Delta x$$

and the total absolute scaled area A^* is:

$$A^*(\eta) = \int_{-\infty}^{\infty} P_p(x)\,P_i(\eta - x)\,dx$$

The reason we say "absolute scaled" is that what is needed is not the scaled area, but the *relative* scaled area. This is understood by looking at the one dimensional uniform densities in Figure 1. Each little strip of P_p that falls under P_i should not be weighted by $1/M$. If we did, we would would end up with a value for the scaled area of $1/N$. That is, the scale factor would have been used to compute the average value of P_p over P_i. The problem is that the desired relative area A has been scaled and we need to normalize by the average scale factor. The average scale factor $\bar{s}$ is the expected value of P_i taken over P_i:

$$\bar{s} = \int_{-\infty}^{\infty} P_i(x)\cdot P_i(x)dx = \int_{-\infty}^{\infty} (P_i(x))^2 dx$$

. This yields the expression for the relative weighted area when the measurement is η:

$$A(\eta) = \frac{A^*(\eta)}{\bar{s}} = \frac{\int_{-\infty}^{\infty} P_p(x)\,P_i(\eta - x)\,dx}{\int_{-\infty}^{\infty}(P_i(x))^2 dx} \tag{3}$$

Applying this formula to our uniform density example gives $\frac{1/N}{1/M} = M/N$ as expected. To compute the expected value of the confusion we need to take the expectation over the population:

$$E[A(x)] = \int_{-\infty}^{\infty} A(x)P_p(x)dx \tag{4}$$

We can analytically apply this measure of expected confusion to the Gaussian case as well. We derive the expression for the one dimensional case; the multi-dimensional case follows naturally. Assume that $P_p(x)$ and $P_i(x)$ are normal densities $N(\mu_p, \sigma_p^2)$ and $N(\mu_i, \sigma_i^2)$ respectively where $\sigma_p \gg \sigma_i$, and μ_i is the mean of the individual. Again considering a measurement of η, the value for A^* is:

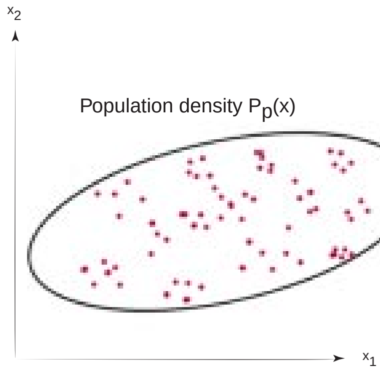
$$A^*(\eta) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_p} e^{\frac{(x-\mu_p)^2}{2\sigma_p^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{(x-\eta)^2}{2\sigma_i^2}}\,dx$$

Given that $\sigma_p \gg \sigma_i$ we can assume that over the region that P_i is significantly non-zero, the value of P_p is constant, in this case namely $P_p(\eta)$. If so, then:

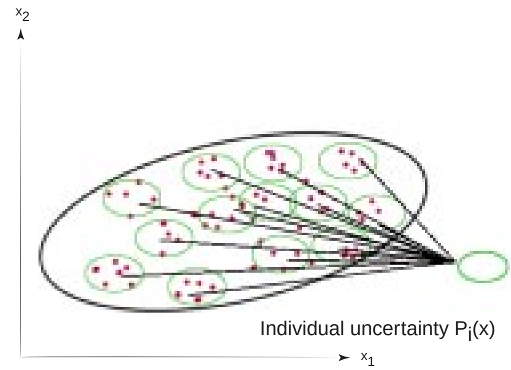
$$\begin{aligned} A^*(\eta) &= P_p(\eta)\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{(x-\eta)^2}{2\sigma_i^2}}\,dx \\ &= P_p(\eta) \end{aligned}$$

The average scale factor $\bar{s}$ is:

$$\bar{s} = \int_{-\infty}^{\infty} \left[\frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{(x-\eta)^2}{2\sigma_i^2}}\right]^2 dx$$



(a)



(b)

Figure 2: This figure shows the multidimensional Gausssian case of the expected confusion. (a) The population density is estimated with samples from a population. (b) The individual uncertainty estimate is based on the average variation of the feature obtain from an individual.

Using a change of variable of $\gamma_i^2 = \sigma_i^2/2$ yields:

$$\begin{aligned} \bar{s} &= \frac{1}{\sqrt{2\pi}\sigma_i \cdot \sqrt{2}} \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\gamma_i} e^{\frac{(x-\eta)^2}{2\gamma_i^2}} dx \\ &= \frac{1}{2\sqrt{\pi}\sigma_i} \end{aligned} \quad (5)$$

This yields a confusion at η for the Gaussian case of:

$$\begin{aligned} A(\eta) &= \frac{A^*(\eta)}{\bar{s}} = \frac{P_p(\eta)}{\frac{1}{2\sqrt{\pi}\sigma_i}} \\ &= 2\sqrt{\pi}\sigma_i P_p(\eta) \end{aligned}$$

Finally to compute the expected confusion we integrate over P_p:

$$\begin{aligned} E[A] &= \int_{-\infty}^{\infty} 2\sqrt{\pi}\sigma_i P_p(\eta) \cdot P_p(\eta) d\eta \\ &= = 2\sqrt{\pi}\sigma_i \int_{-\infty}^{\infty} (P_p(\eta))^2 d\eta \\ &= 2\sqrt{\pi}\sigma_i \cdot \frac{1}{2\sqrt{\pi}\sigma_p} \\ &= \sigma_i/\sigma_p \end{aligned} \quad (6)$$

This satisfying result simply states that the percentage of the population remaining is the ratio of standard deviation of the uncertainty after measurement to that before the measurement is taken. If the $-\ln$ of the ratio is taken,

$$-\ln(\frac{\sigma_i}{\sigma_p}) = \ln\sigma_p - \ln\sigma_i, \quad (7)$$

we arrive at an expression that is the mutual information from Equation 2. This is shown by solving Equation 2 using the random variables $\mathbf{X}$ (the population density) and $\mathbf{Y}$ (the individual uncertainty) defined by their respective Gaussian distributions:

$$P_p(x) = \frac{1}{\sqrt{2\pi}\sigma_p} e^{\frac{x^2}{2\sigma_p^2}}$$

and

$$P_i(y) = \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{y^2}{2\sigma_i^2}}.$$

First, solving for H($\mathbf{X}$), we arrive at:

$$\begin{aligned} H(\mathbf{X}) &= -\int_x P_p(x) \ln P_p(x) dx \\ &= -\int_x P_p(x)[\frac{-x^2}{2\sigma_p^2} - \ln\sqrt{2\pi\sigma_p^2}]dx \\ &= \int_x \frac{x^2 P_p(x)}{2\sigma_p^2} dx + \int_x P_p(x) \ln\sqrt{2\pi\sigma_p^2} dx \\ &= \frac{1}{2\sigma_p^2} \int_x x^2 P_p(x) dx + \ln 2\pi\sigma_p^2 \int_x P_p(x) dx \\ &= \frac{\sigma_p^2}{2\sigma_p^2} + \frac{1}{2}\ln 2\pi\sigma_i^2 \\ &= \frac{1}{2} + \frac{1}{2}\ln 2\pi\sigma_p^2 \end{aligned}$$

Second, solving for H($\mathbf{Y}$), we arrive at a similar expression:

$$\frac{1}{2} + \frac{1}{2}\ln 2\pi\sigma_i^2$$

Last, solving for $I(\mathbf{X};\mathbf{Y})$:

$$
\begin{aligned}
I(\mathbf{X};\mathbf{Y}) &\equiv H(\mathbf{X}) - H(\mathbf{Y}) \\
&\equiv \frac{1}{2} + \frac{1}{2}\ln 2\pi\sigma_p^2 - \frac{1}{2} - \frac{1}{2}\ln 2\pi\sigma_i^2 \\
&\equiv \frac{1}{2}\ln(\frac{2\pi\sigma_p^2}{2\pi\sigma_i^2}) \\
&\equiv \ln\sigma_p - \ln\sigma_i. \qquad (8)
\end{aligned}
$$

The result (Equation 8) is equal to Equation 7, and shows the relationship of our expected confusion metric and mutual information. However, our measure is a relative one and can more easily be interpreted as a percentage of overlap and thus is perhaps more directly applicable to the field of biometrics.

Extending the expected confusion measure to the multi-dimensional Gaussian case, the result is

$$
\text{Expected Confusion} = \frac{|\Sigma_i|^{1/2}}{|\Sigma_p|^{1/2}}. \qquad (9)
$$

This quantity is the ratio of the volumes of equal probability hyper-ellipsoids as defined by the Gaussian densities. $|\Sigma_p|^{1/2}$ is computed from the population density (see Figure 2(a)), and $|\Sigma_i|^{1/2}$ is computed from the average individual variation (see Figure 2(b)). This analysis also holds for densities that are mixtures of Gaussians.

3. Static Measurements from Gait

In this section we present a set of static body parameters designed for gait from our pervious work [2] to demonstrated the use of the expected confusion metric. However, unlike [2], we recover the static body parameters from a motion-capture database of walking subjects instead of a vision database.

3.1. Gait parameters

Our motion-capture system uses magnetic sensors to capture the three-dimensional position and orientation of the limbs of the subjects as they walk along a platform. Sixteen sensors in all are used for the head, torso, pelvis, hands, forearms, upper-arms, thighs, calves, and feet. For this experiment, we recorded 20 subjects (11 males and 9 females with heights varying from 149.9 cm to 185.4 cm) walking along a 16ft. long platform.

The static body parameters measured, as a person walks, are four distances: the distance between the head and foot (d_1), the distance between the head and pelvis (d_2),the distance between the foot and pelvis (d_3), and the distance between the left foot and right foot (d_4) (See Figure 3). These distances are only measured at the maximal separation point of the feet during the double support phase of the gait cycle, and are concatenated to form a four-dimensional walk vector $\mathbf{w} =< d_1, d_2, d_3, d_3 >$ for each subject.

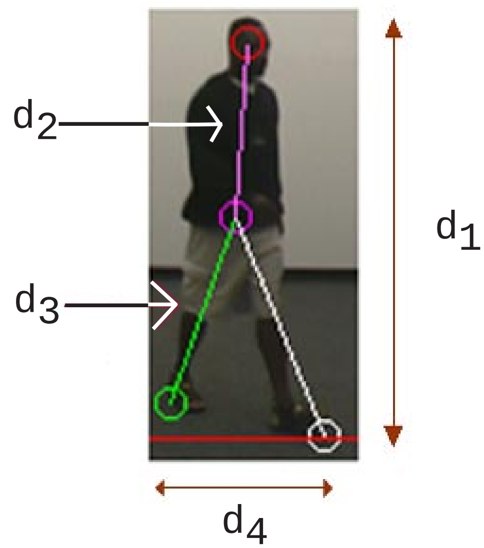


Figure 3: The static body parameters: $\mathbf{w} =< d_1, d_2, d_3, d_4 >$.

3.2. Population vs. individual variation

Following the analysis of Section 2, we want to measure the ratio between the volume of the individual variation density and that of the overall population. Because of a limited number of subjects we will model the density of the population as a single Gaussian. To determine whether our density estimation is valid, we plot the value of $|\Sigma_p|^{\frac{1}{2}}$ as we add more subjects (see Figure 4). The data point for k subjects was computed by taking 200 random sets[1] of k subjects and computing the maximum likelihood estimate Gaussian density. The fact that the volume begins to asymptote as we approach 18 subjects implies that we have a reasonable model of the population density, but of course more data is always better. The asymptotic value is 309 cm^4.

To compute the individual variation we subtract the mean walk vector of each subject from each of their six trials, and then compute the covariance Σ_i over all the trials. The value of $|\Sigma_i|^{\frac{1}{2}}$ for the first set of static body parameter is 1.3 cm^4. This yields an expected confusion ratio

$$
E[A] = \frac{1.3cm^4}{309cm^4} = 0.0042.
$$

This implies that the variation in these static body parameters during an individual's gait would leave a confusion with an average of less than 1% of the population. If the expected confusion ratio was equal to one, then the individual variations of the static body parameters would be the same as

[1]In cases where $\begin{pmatrix} n \\ k \end{pmatrix}$ is less than 200, all possible sets of k were used.

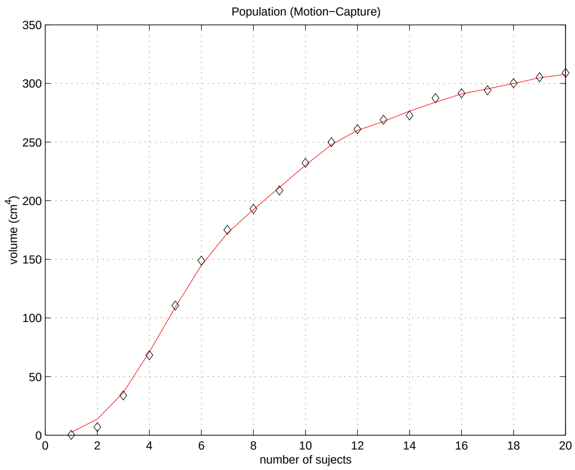


Figure 4: Volumetric measure of magnitude of probability density $|\Sigma_p|^{\frac{1}{2}}$ of the population using the motion-capture data for walk vector $\mathbf{w}$. The curve reaches a somewhat stable asymptote after approximately 18 subjects indicating a reasonable coverage of (a segment of) the population.

the population variation of the parameters, and would imply that the features are not discrimination between individuals.

We note that when performing recognition on our database of 20 people, our recognition rate was 97.5%. But clearly this number is a function of how many people are in our database. The confusion measure, however, is valid as soon as the number of subjects allows the estimate of the population density to converge.

4. Conclusion

This paper has given a derivation of an expected confusion measure as opposed to reporting a percent correct with a limited database. Even with only 20 subjects we are able to make predictions as to how well a particular measure will perform. As with any new work, there are several next steps to be undertaken. We must expand our database to test how well the expected confusion metric predicts performance over larger databases. It will also permit considering whether representing the population as a single Gaussian density is wise.

References

[1] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. John Wilety & Sons, Inc., New York, 1991.

[2] A. Y. Johnson and A. F. Bobick. A multi-view method for gait recogntion using static body parameters. *3rd International Conference on AUDIO- and VIDEO-BASED BIOMETRIC PERSON AUTHENTICATION*, Halmstad, Sweden, June 6-8, 2001.