Available online at www.sciencedirect.com

ScienceDirect

Forensic Science International 177 (2008) 228–237

Forensic Science International

www.elsevier.com/locate/forsciint

Obtaining confidence intervals and Likelihood Ratios for body height estimations in images

Ivo Alberink *, Annabel Bolck

Netherlands Forensic Institute, Department of Digital Technology and Biometry, P.O. Box 24044, 2490 AA, Den Haag, The Netherlands

Received 23 August 2007; received in revised form 20 December 2007; accepted 9 January 2008
Available online 3 March 2008

Abstract

In forensic practice, height estimations on perpetrators visible in video footage from surveillance cameras are regularly requested. There are several ways to do this. Insight is gained into the difference between actual and measured heights by taking validation measurements of a number of test persons. Variation between actual and measured heights is decomposed into a systematic part (because of height loss by pose, 3D modeling of the scene of crime, operator biases) and a random part (due to natural variation). On this basis a method is described for obtaining confidence intervals for the height, including head- and footwear, of questioned persons in images. Since the number of test persons is usually limited, the result is in terms of the Student's t distribution. In addition, for cases in which a suspect is available, an expression is obtained for the Likelihood Ratio (LR), measuring the strength of evidence of resemblance of actual height of the suspect and measured height of the perpetrator. The Likelihood Ratio depends both on the rarity of the estimated perpetrator's height and its closeness to the suspect's height. Technical theorems included may be relevant for other forensic areas as well.

Keywords: Body height estimation; Confidence intervals; Likelihood Ratios

1. Introduction

In forensic practice, height estimations on perpetrators visible in video footage from surveillance cameras are regularly requested. The approach to this at the Netherlands Forensic Institute is the following.

The scene of crime is visited with a number of test persons. Perpetrator and test persons are referred to as *donors*. A Closed Circuit Television (CCTV) or camera image is selected in which the perpetrator is standing more or less in upright position. The test persons are positioned at the same location and in front of the same camera as the perpetrator on the original footage, in as much as possible the same pose. This procedure is called a *reconstruction* and yields validation readings that allow to interpret height estimations on the perpetrator correctly.

Moreover, on the basis of 2D photographs and fixed location points, a 3D model of the scene of crime is created. Using common points of the 3D model and the camera view on the questioned image, the location and orientation of the camera is determined, and the 3D model is projected such that it has the same perspective as the camera images. Next, up to four investigators, or *operators*, perform height measurements on the donors by placing cylinders over the bodies in the 3D model, from feet to head. The height of the cylinders approximates the actual height of the donors, reduced by the loss in height by the pose of the perpetrator. To reduce variation due to the operators, the procedure is repeated for, say, three times, resulting in 3×4 height measurements per donor, or four mean height measurements, one for each operator.

Variation between actual and measured heights of donors is introduced by the following factors:

1. Creation of the 3D model,
2. Finding of camera position, orientation and focal length,
3. Presence of lens distortion at the location of the perpetrator in the chosen image,
4. Pose of the perpetrator in the chosen image,
5. Presence and height of head- and footwear,
6. Interpretation of head and feet in the images by the operators.

* Corresponding author. Tel.: +31 70 888 6400; fax: +31 70 888 6559.
E-mail address: i.alberink@nfi.minjus.nl (I. Alberink).

0379-0738/$ – see front matter
doi:10.1016/j.forsciint.2008.01.005

This variation may be decomposed into a

1. Systematic part, because of
   - modeling of the scene of crime (points 1, 2, 3),
   - height loss because the perpetrator (and hence the rest of the donors) does not stand up straight (4).
2. Random part, because of varying head- and footwear and operators doing measurements that have natural variation (5 and 6).

By measuring reference objects in the image, like measuring sticks, an estimate of the systematic error by variation in the modeling of the crime scene can be made. Systematic error by varying height loss because of pose cannot be estimated directly. In practice (casework) systematic errors amount to several centimeters and vary from case to case. Since variation introduced by head- and footwear cannot be removed without extra knowledge, height measurements are usually of donors including head- and footwear.

The paper aims at answering the following two questions:

1. On the basis of the measurements, how can we give probability statements (confidence intervals) on the actual height of the perpetrator?
2. In case there is a suspect: what is the evidential value, in terms of a Likelihood Ratio (LR), of eventual resemblance of suspect's and perpetrator's height?

As described in [1], these are the key questions in the forensic height measuring process. The questions did not receive much attention in literature though, which has been focusing more on technical methods than validation.

For literature on body height estimations in digital images, cf. [1–6]. In [7], the situation is considered of images from separate crime scenes that may or may not contain the same perpetrator. A theoretical explanation is given on how to determine the LR connected to resemblance in body height of the persons in the images. In [8], an LR is determined if an eyewitness statement about the body height of a perpetrator concurs with the height of a suspect. Both papers [7,8] do not concentrate on the problem of the limited number of test persons used though, which seems to be the central problem in casework. The current paper aims to give a statistically sound approach in this respect.

The layout of the paper is as follows. Section 2 contains the methods and techniques used, describing notation, height measurement method, the central numerical example and the statistical model. In Section 3 it is described how confidence intervals for the perpetrator's height may be constructed. Section 4 is about obtaining Likelihood Ratios for comparing perpetrators and suspects. In Section 5 conclusions on and a discussion of the results are given.

2. Methods and techniques

We describe the technical and mathematical methods used to answer the questions. Section 2.1 introduces general notation on normal distributions and the taking of averages of measurements. In Section 2.2, more detail is given about the method used for performing the height measurements. Section 2.3 describes the case example along the lines of which all results will be presented. Finally we describe the statistical model underlying the whole analysis.

2.1. Notation

From here on, for a random variable X which is normally distributed with mean μ and variance σ^2, we denote $X \sim N(\mu,\sigma^2)$. In the multivariate case, if a random vector $\mathbf{X}$ has the Gaussian (normal) distribution with mean vector $\boldsymbol{\mu}$ and covariance matrix $\mathbf{C}$, see e.g. [9], Chapter 8, this is denoted by $\mathbf{X} \sim N(\boldsymbol{\mu},\mathbf{C})$. The univariate normal probability density function is given by

$$\varphi_{\mu,\sigma}(x) = (2\pi)^{-1/2}\,\sigma^{-1}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right).$$

Moreover, for any independent sample of random variables $X_1, \ldots, X_n$ we will denote the sample mean and sample variance by

$$\bar{X} = \frac{1}{n}\sum_{j=1}^{n} X_j \quad \text{and} \quad S_X^2 = \frac{1}{n-1}\sum_{j=1}^{n}(X_j - \bar{X})^2.$$

In case of multiple subscripts, e.g. for the sequence

$$X_{k_1,k_2,k_3} : 1 \le k_1 \le n_1, 1 \le k_2 \le n_2, 1 \le k_3 \le n_3,$$

bullets (•) are used to indicate with respect to which subscript an average is taken, that is

$$X_{k_1,\bullet,k_3} = \frac{1}{n_2}\sum_{k_2=1}^{n_2} X_{k_1,k_2,k_3}, \qquad X_{k_1,\bullet,\bullet} = \frac{1}{n_2 n_3}\sum_{k_2=1}^{n_2}\sum_{k_3=1}^{n_3} X_{k_1,k_2,k_3},$$

etc.

2.2. Method used for performing height measurements

The method used at the Netherlands Forensic Institute for doing height measurements in images is through the construction of a 3D model of the crime scene. This can be done in different ways. An often-used technique is by means of photogrammetric software. This software uses the fact that 3D coordinates of fixed location points can be measured by making photos of the scene from different positions. When enough common points are identified on each image a 3D model of the scene is made, consisting of the set of points. (Instead of using photographs, a laser scanner can be used for the construction of a 3D model of the scene.)

Next, a human operator links points of the 3D model to corresponding points in the questioned image. This makes it possible to determine position, rotation and focal length of the camera taking the images. The procedure of finding the right camera parameters is referred to as doing a *camera match*. Using the camera information, a virtual camera can then be placed in the 3D model of the room, looking at the model from

Fig. 1. Example of a 3D computer model and its projection on the image (bottom right).

the same perspective as the real camera at the real crime scene. Fig. 1 shows a screen shot of the software 3dsMax on which a 3D model of a simulated crime scene can be seen, together with its projection on a questioned image.

On the basis of the 3D model and camera parameters, in software like 3dsMax it is possible to measure heights and distances on the image. The height of a person can be measured by placing a cylinder in the scene, from feet to the top of the head, or by creating human models (so-called bipeds), that can be adjusted to the stance of the person. The process of doing measurements is depicted in Fig. 2.

An advantage of the method is the wide applicability in forensic casework, since a 3D model of the crime scene can always be made, independent of the surveillance camera and the available images. A disadvantage is that taking the pictures is time consuming and thus expensive.

Alternative methods include single view metrology on the basis of vanishing points, cf. [2,5]. A review of the methods is given in [1].

2.3. Case example

We present the analysis along the lines of the following case example. In Fig. 3, video footage is shown of a robbery of a cinema.

A reconstruction was made on the crime scene, positioning six test persons of various heights in front of the camera that yielded the image of the perpetrator. A 3D model of the room

Fig. 2. Example of performance of a height measurement using either a cylinder (on the left) or a biped (on the right).

Fig. 3. Case example: questioned image. A height measurement was requested of the person in the oval.

was made and four operators performed height measurements on chosen images of all donors. The operators repeated the procedure for three times, on separate occasions, with the images in randomized order. The averaged measured heights per operator are described in Table 1.

The concept *actual height* of donors will be defined as follows: it is the average over three measurements (by different operators) on the donor, using a ruler like the one shown in Fig. 4.

Actual heights of the test persons next to the means over all measured heights per donor (including head- and footwear, and if known) are given in Table 2.

In Section 3, it is described how to obtain a confidence interval of the actual height of the perpetrator from the above. In Section 4, we describe how to obtain a Likelihood Ratio for interpreting evidential value of the results against a suspect.

2.4. Statistical model

We define the statistical model underlying the analysis. For any donor j, let the actual height L_j be the average over three measurements by different operators, using a ruler like the one shown in Fig. 4. Moreover, let $M_{j,k,l}$ be the measured height of donor j, measured by operator k, repeated measurement l.

Table 1
Measured heights for all donors, in cm, averaged over each operator's measurements

Operator	1	2	3	4
Perpetrator	163.7	167.1	166.9	168.0
Test person				
1	185.0	187.4	185.5	188.3
2	164.3	169.1	165.8	167.0
3	157.7	159.0	159.0	160.7
4	173.0	176.9	175.8	177.0
5	177.3	180.4	178.6	180.3
6	185.3	188.5	185.5	189.3

As explained in the introduction, the difference between actual and measured height is influenced by:

- The height loss by the pose of the perpetrator.
- The 3D model of the crime scene.
- The camera match that is used.
- Random variation because of natural variation in the results of operators.

In casework, loss in height because of pose is more or less uniform since all donors are positioned in a way similar to the perpetrator. Slight deviations will occur, because positioning is never perfect, but these may be considered as random (not systematic), per person. The 3D model of the crime scene is the same for every donor and hence does not lead to different readings for different donors.

With respect to the camera match: it is specific to the individual operators, that is to say, each operator uses his own camera match for all donor images. This may cause an operator specific bias in the observed differences between actual and measured height. Apart from the above, in [6] an operator specific component in height measurements is reported that does not depend on the camera match alone. Note however that the biases are constant for all donors so that operator effects will even out.

We want to estimate the perpetrator's actual height L_p given the measured heights $M_{\mathrm{p},k,l}$. Hence, natural interest lies at gaining insight about differences, per person, between actual and measured heights. Following the notation of Section 2.1, let $M_j = M_{j,\bullet,\bullet}$ be the mean of all height measurements on donor j. We analyze the 'differences':

$$\Delta_j = L_j - M_{j,\bullet,\bullet}$$

between actual and measured heights. If we follow the assumptions stated above, the differences are mutually independent, with

$$\Delta_j \sim N(\mu_{\mathrm{pos}} + \mu_{\mathrm{opt},\bullet}, \sigma^2_{\mathrm{pos}} + \sigma^2_{\mathrm{opt},\bullet}),$$

with unknown means and variances according to positioning ('pos') and operators ('opt'), respectively. Note that in the model, variation by height loss and by operator 'inaccuracy' add up.

The bottom line of the above is the following. We assume that the averaged differences between actual and measured heights, per donor, are independent, identically distributed variables with unknown mean and variance. On the basis of the validation experiment the unknown mean and variance are studied, and using Theorem A1 in Appendix A.1 this leads to a probability distribution on any new difference, especially for that of the perpetrator. The latter makes it possible to calculate a confidence band for the perpetrator's height.

3. Confidence intervals

We describe how to determine confidence intervals for the height of the perpetrator, including head- and footwear.

The general result is described in Theorem 1, following the notation of Section 2.4. That is to say, let n be the number of test

Fig. 4. Ruler used to measure the actual height of test persons and suspect.

persons used in the validation experiment, let Δ_j be the averaged difference between actual and measured height per donor, and let $M_p = M_{p,\bullet,\bullet}$ be the average measured height of the perpetrator. Let $\bar{\Delta}$ and S^2_Δ be the sample mean and the sample variance of the differences.

Theorem 1. *Let* $0 < \alpha < 1$, *and let* $t_{n-1,1-\alpha/2}$ *be the* $1-\alpha/2$ *percentile of the Student's t distribution with* $n-1$ *degrees of freedom. Then given* M_p *and the differences* Δ_j, *a* $100\,(1-\alpha)\%$ *confidence interval for* L_p *is given through*:

$$|L_p - (M_p + \bar{\Delta})| \le (1+n^{-1})^{1/2} S_\Delta\, t_{n-1,1-\alpha/2}.$$

Proof. According to Theorem A1 of Appendix A.1, with respect to the mean total error $\Delta_p(=L_p - M_p)$ of the perpetrator, the variable

$$V = \frac{\Delta_p - \bar{\Delta}}{(1+n^{-1})^{1/2} S_\Delta}$$

has the Student's t distribution with $n-1$ degrees of freedom. Hence indeed

$$P(|\Delta_p - \bar{\Delta}| \le (1+n^{-1})^{1/2} S_\Delta\, t_{n-1,1-\alpha/2}) = 1-\alpha. \qquad \square$$

In words: given the mean measured height of the perpetrator and the estimated mean systematic bias, a $100 \times (1-\alpha)\%$ confidence interval for the actual height of the perpetrator is given by the sum of the mean measured height of the perpetrator and the estimated mean systematic bias ($M_p + \bar{\Delta}$), plus and minus a constant ($(1+n^{-1})^{1/2}$) depending on the sample size, times the estimated standard deviation (S_Δ) of the mean biases found, times the $1-\alpha/2$ percentile of a Student's t distribution depending on the sample size ($t_{n-1,1-\alpha/2}$).

In the numerical example this means the following. Based on the outcomes in Table 1, the mean measured height of the perpetrator is 1.664 m. For the test persons the mean measured height is systematically 6.3 cm lower than actual height (including head- and footwear), with a standard deviation of 1.7 cm. Hence an approximation of the actual height of the perpetrator (including head- and footwear) is given by 166.4 + 6.3 = 172.7 cm with some confidence band. The width of the band is determined by application of Theorem 1. In the terminology of Section 2.1:

$$\bar{\Delta} = 6.3\,\text{cm} \quad \text{and} \quad S_\Delta = 1.7\,\text{cm},$$

where $\bar{\Delta}$ and S_Δ are the mean and standard deviation of the numbers in column 4 of Table 2, whereas $t_{6-1,1-0.025} = t_{5,0.975} = 2.571$, see Appendix B. As a result, with 95% confidence:

$$|L_p - (M_p + \bar{\Delta})| \le 2.571(1+6^{-1})^{1/2} S_\Delta = 4.7\,\text{cm}.$$

Table 2
Actual heights, mean measured heights, and their differences for all donors, plus the actual height of the suspect, in cm

	Actual height	Mean measured height	Difference
Perpetrator	?	166.4	?
Test person			
1	193.8	186.5	7.3
2	174.3	166.6	7.7
3	163.0	159.1	3.9
4	181.9	175.7	6.2
5	187.3	179.2	8.1
6	191.9	187.2	4.7
Suspect	176.2	–	–

This means that the actual height of the perpetrator, including head- and footwear, with 95% confidence is contained in the interval

$$(172.7 - 4.7; 172.7 + 4.7\,\text{cm}) = (168;\ 177.5\,\text{cm}).$$

Considering that the suspect has an actual height, including head- and footwear, of 176 cm, the hypothesis that suspect and perpetrator are the same person is not rejected (with a significance level of 95%).

4. Evidential value: Likelihood Ratios

We turn to the concept of evidential value of the resemblance between the estimated height of the perpetrator and the actual height of a suspect. In Section 4.1, a general description of the concept Likelihood Ratio is given. In Section 4.2, we discuss how to optimally choose the variable that is evaluated in the definition of the LR. In Section 4.3 necessary assumptions are stated. In Sections 4.4 and 4.5, the main result on Likelihood Ratios (Theorem 2) and an evaluation of the result are described. In Section 4.6 an alternative is given on the assumption on head- and footwear.

4.1. Likelihood Ratios

We test the 'prosecutor's hypothesis'

H_0: Suspect and perpetrator are identical
against the 'defense hypothesis'

H_1: Suspect and perpetrator are not identical.

The best way of evaluating the evidence is by means of the so-called *Likelihood Ratio*—from here on: LR. For discrete evidence (like DNA profiles), a usual way to define the LR is by the formula:

$$\text{LR} = \frac{P(\text{Evidence}|\text{H}_0)}{P(\text{Evidence}|\text{H}_1)}.$$

In this way, the LR expresses in how far the evidence was more probable under H_0 than under H_1, and hence is a numerical representation of the strength of the evidence. For a rigorous explanation of the use of LRs in forensic practice, see [8,10].

When looking at scores of fingerprint comparisons, comparison of concentrations of chemical components, or in the present case at height measurements, the evidence consists of continuous readings. The evidence is then first summarized into a relevant (uni- or multivariate) variable T, consisting of the evidence itself or a derivation from it, for example, all reference and control samples, or just their means. Given that $T = t$, the LR for this outcome is then defined as

$$\text{LR}(t) = \text{LR}_T(t) = \frac{f_T^0(t)}{f_T^1(t)} \qquad (*)$$

the ratio of the *probability densities* of T in t given H_0 and H_1, respectively. The same function T of the evidence has to be evaluated in the numerator and denominator of the definition of the LR.

An example of this is the following. Suppose glass is recovered from a crime scene for which the refractive index is measured a number of times, leading to reference measurements $X = (X_1, \ldots, X_m)$. Furthermore measurements are taken on a suspect piece of glass, giving control measurements $Y = (Y_1, \ldots, Y_n)$. The strength of evidence for the hypothesis that the suspect glass came from the glass at the scene of crime (H_0) is then quantified in the formula:

$$\text{LR}(x_1, \ldots, x_m, y_1, \ldots, y_n) = \text{LR}(x, y) = \frac{f_{X,Y}^0(x, y)}{f_{X,Y}^1(x, y)},$$

where $f_{X,Y}^0(x, y)$ and $f_{X,Y}^1(x, y)$ are joint density functions of X and Y under the prosecutor's and defense hypothesis, respectively. Under assumptions of normality and knowledge about refractive indices on the 'glass population' one may express the LR in a closed formula, cf. [10], Section 10.7.

4.2. Choice of appropriate statistic

Let as before L_j, M_j and Δ_j denote actual and measured height of person j, including head- and footwear, and the differences between them. The 'evidence', that is to say, the relevant information gathered with respect to resemblance of height of perpetrator and suspect, may be described as

$$\text{Evidence} = (M_\text{p}, L_\text{s}, \Delta_1, \ldots, \Delta_n).$$

Note that under H_0, $L_\text{s} \approx M_\text{p} + \bar{\Delta}$, whereas under H_1 the two (L_s and $M_\text{p} + \bar{\Delta}$) are statistically independent. Possible choices for the LR test statistic T are:

- $T_1 = (L_\text{s}, M_\text{p})$,
- $T_2 = (L_\text{s} - (M_\text{p} + \bar{\Delta}))/S_\Delta$.

Since the probability distribution of L_s given M_p is unknown under H_0 we cannot evaluate the numerator of (*) for T_1.

Using T_2 is tempting under H_0 since then T has a Student's t distribution (up to a constant) which removes the problem of estimating the unknown variance of the differences Δ_j. However, under H_1 the normalization by S_Δ gives a complicated probability distribution for T_2, making it an inconvenient choice as well.

Therefore, we leave S_Δ out and evaluate

$$T = (L_\text{s}, M_\text{p} + \bar{\Delta}).$$

Based on knowledge of the height distribution, excluding head- and footwear, of the target population and on the error measurements $\Delta_1, \ldots, \Delta_n$ we calculate numerator and denominator of (*).

4.3. Assumptions

In order to calculate numerator and denominator of the LR, we need to do some model assumptions for population and error measurements:

Assumption 1. The height distribution for the specific population that is looked at, in this case Dutch Caucasian males, has a $N(\mu,\tau^2)$ distribution.

Assumption 2. Differences Δ_j are normally distributed and statistically independent of the heights L_j of the donors.

If it is clear which (sub)population to use, Assumption 1 is usually well founded and the mean and variance are known. With respect to Assumption 2 we remark that in [6] of the three security cameras that were used, no significant influence of actual height of persons on measurement errors was found on two of the three. (The third camera, located at a height of ca. 6 m and with a rather low resolution of people in the images, showed mixed results in this respect.)

Concerning head- and footwear, we need to consider the following:

1. The height distribution of the population does not include head- and footwear, whereas the measurements of actual and measured heights do.
2. Head- and footwear of the perpetrator are not the same as that of the suspect.

We initially handle this by using the following assumption:

Assumption 3. For any person in the population, a constant height of ν cm should be added to counter for the height difference caused by head- and footwear.

On the basis of Assumptions 1 up to 3 a first analysis on LR calculation can be performed.

4.4. Calculation of the LR

First we formulate the main result. Let as before L_s be the actual height of the suspect, and $M_p = M_{p,\cdot,\cdot}$ the mean measured height of the perpetrator in the image. Let moreover $\Delta_1, \ldots, \Delta_n$ denote differences between actual and measured heights for n test persons, with average $\bar{\Delta}$. Assuming that Assumptions 1 up to 3 hold, that is

1. the height distribution over the population is $N(\mu,\tau^2)$-distributed,
2. the Δ_j are normally distributed, with variance σ^2,
3. head- and footwear add a constant ν cm to a person's height,

the LR reads as follows:

Theorem 2. *Under Assumptions 1 up to 3, we have*:

$$\mathrm{LR} = \sqrt{\frac{\tau^2+(1+n^{-1})\sigma^2}{(1+n^{-1})\sigma^2}}\exp\left(\frac{1}{2}\left(\frac{(M_p+\bar{\Delta}-(\mu+\nu))^2}{\tau^2+(1+n^{-1})\sigma^2}-\frac{(L_s-(M_p+\bar{\Delta}))^2}{(1+n^{-1})\sigma^2}\right)\right).$$

Proof. Following the definition of the LR, we evaluate its numerator and denominator

$$f_T^0(t) = f_T^0(l_s, m_p+\bar{\delta}) \quad \text{and} \quad f_T^1(t) = f_T^1(l_s, m_p+\bar{\delta}),$$

with $T = (L_s, M_p + \bar{\Delta})$.

We start by looking at the numerator. Under the prosecutor's hypothesis the components of T are highly correlated, since then $L_s = L_p$ and thus $T = (L_p, M_p + \bar{\Delta})$. For the pair of variables $(L_p, \Delta_p - \bar{\Delta})$ however, following Assumption 2 it is clear that they are statistically independent and hence much easier to handle. Using Theorem A2 of Appendix A.2 we have

$$f_T^0(l_s, m_p+\bar{\delta}) = f_{AT}^0(l_s, l_s - m_p - \bar{\delta})$$

with

$$A = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}.$$

As stated the components of AT are independent and normally distributed, with

$$L_p \sim N(\mu,\tau^2) \quad \text{and} \quad \Delta_p - \bar{\Delta} \sim N(0, (1+n^{-1})\sigma^2),$$

and we see that

$$f_T^0(t) = f_{AT}^0(l_s, l_s - m_p - \bar{\delta}) = \varphi_{\mu+\nu,\tau}(l_s)\varphi_{0,(1+n^{-1})^{1/2}\sigma}(l_s - m_p - \bar{\delta}).$$

Turning to the denominator: under H_d the components of T are mutually independent, so that

$$f_T^1(t) = f_{L_s}(l_s)\,f_{M_p+\bar{\Delta}}(m_p+\bar{\delta}) = \varphi_{\mu+\nu,\tau}(l_s)\,f_{L_p+(\bar{\Delta}-\Delta_p)}(m_p+\bar{\delta}).$$

Here L_p and $\bar{\Delta} - \Delta_p$ are independent and normally distributed, and hence their sum is normally distributed as well, with mean $\mu + \nu$ (Assumption 3) and variance $\tau^2 + (1 + n^{-1})\,\sigma^2$. As a consequence

$$f_T^1(t) = \varphi_{\mu+\nu,\tau}(l_s)\varphi_{\mu+\nu,(\tau^2+(1+n^{-1})\sigma^2)^{1/2}}(m_p+\bar{\delta})$$

and the result of the theorem easily follows. □

4.5. Numerical examples

In the numerical example:

$$n = 6, M_p = 166.4\,\text{cm}, L_s = 176.2\,\text{cm}, \quad \bar{\Delta} = 6.3\,\text{cm and } S_\Delta = 1.7\,\text{cm}.$$

According to the Dutch Central Bureau of Statistics, for the population of Dutch Caucasian males, mean and standard deviation of the population (anno 2006) are given by $\mu_{NL} = 180.6$ cm and $\tau_{NL} \approx 10$ cm. Assuming furthermore that $\nu = 2$ cm and $\sigma \approx S_\Delta$, Theorem 2 yields:

$$\mathrm{LR} = \sqrt{\frac{\tau^2+(1+n^{-1})\sigma^2}{(1+n^{-1})\sigma^2}}\exp\left(\frac{1}{2}\left(\frac{(M_p+\bar{\Delta}-(\mu+\nu))^2}{\tau^2+(1+n^{-1})\sigma^2}-\frac{(L_s-(M_p+\bar{\Delta}))^2}{(1+n^{-1})\sigma^2}\right)\right) \approx 2.$$

This means that the height measurements on the perpetrator are two times more likely when the suspect and perpetrator are identical than when they are not. This constitutes very weak evidence against the suspect.

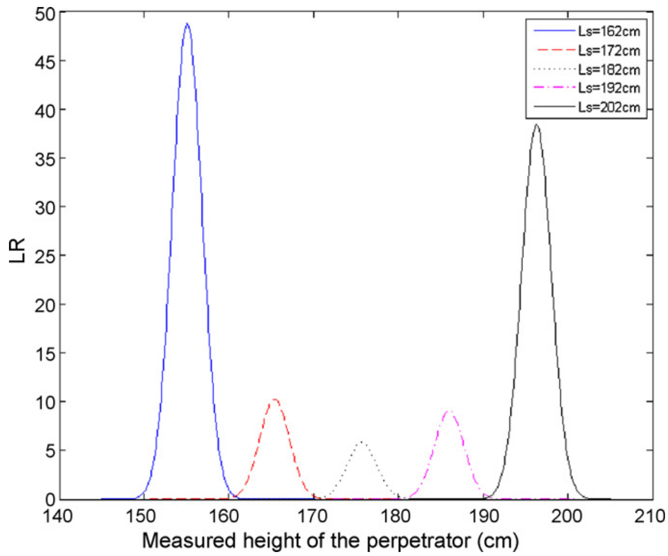

Fig. 5. Examples of the LR function for different suspect heights.

To obtain an impression on LR outcomes that are obtainable, we plot the LR as a function of the measured height of the perpetrator (the evidence), starting from separate fixed suspect heights. We restrict the number of parameters by taking $n = 6$, $\bar{\Delta} = 6.3$ cm, $\mu = 180.6$ cm, $\nu = 2$ cm and $\tau = 10$ cm as above. These values for the parameters only shift the plots ($\bar{\Delta}$, μ, ν) or are more or less fixed (n usually equals 4, 5 or 6, τ is usually close to 10 cm). Furthermore we fixate measurement accuracy by taking $\sigma = 1.7$ cm. We let the suspect's height vary over the population, taking $L_s = 162, 172, 182, 192, 202$ cm and plot the functions $\mathrm{LR} = \mathrm{LR}(M_p)$ in Fig. 5.

We see that for common suspect heights, the obtainable LR reaches up to 6, whereas for example for $L_s = 162$ cm, it reaches up to 50.

4.6. Interpretation of the main result on the Likelihood Ratio

We meditate on the result of Theorem 2. If the target population consists for example of male Dutch Caucasians (or British, cf. [8], Section 12.5), typically $\tau \approx 10$ cm. Moreover, in casework the number of test persons is usually limited, say $n = 4, 5, 6$. We have a look at the formula:

$$\mathrm{LR} = \sqrt{\frac{\tau^2+(1+n^{-1})\sigma^2}{(1+n^{-1})\sigma^2}}\exp\left(\frac{1}{2}\left(\frac{(M_p+\bar{\Delta}-(\mu+\nu))^2}{\tau^2+(1+n^{-1})\sigma^2}-\frac{(L_s-(M_p+\bar{\Delta}))^2}{(1+n^{-1})\sigma^2}\right)\right).$$

Here

1. The term

$$\frac{(M_p+\bar{\Delta}-(\mu+\nu))^2}{\tau^2+(1+n^{-1})\sigma^2}$$

reflects the rarity of the estimated perpetrator's height in the population. If estimated perpetrator's height is far from the population's mean, relative to the variance of the height of the population, the term increases and the LR increases exponentially.

2. The term

$$\frac{(L_s-(M_p+\bar{\Delta}))^2}{(1+n^{-1})\sigma^2}$$

reflects whether the suspect's height and the estimated perpetrator's height are close to one another. If estimated perpetrator's height is far from the suspect's height, relative to the variance of differences perceived on test persons, the term increases and the LR decreases exponentially.

In short: the expression in Theorem 2 reflects both rarity of the estimated perpetrator's height and closeness of it to suspect's height.

The term $\tau^2 + (1 + n^{-1})\,\sigma^2$ is approximated well by τ^2, which may be used to obtain a simplified formula.

Note that the result is in terms of $\sigma^2 = \sigma_\Delta^2$ instead of S_Δ^2. This is a direct result of the particular choice of T. Though we were able to replace the unknown expectation value μ_Δ by its estimator $\bar{\Delta}$ (by introducing $\bar{\Delta}$ into the definition of T), hence removing the problem of the fact that it has to be estimated, we did not succeed in removing the variance of the differences. Note as well that for confidence intervals we were able to solve this problem.

In the numerical example, we used the approximation $\sigma \approx S_\Delta$. On the basis of S_Δ, one may construct a confidence interval for σ, thus obtaining a confidence interval for the Likelihood Ratio. Another possibility to get around the approximation is to introduce a prior distribution on σ_Δ^2 in the proof of Theorem 2, based on results found in casework, and factor it out. Future research on this topic seems sensible.

4.7. Alternative approach on head- and footwear

An alternative to the above approach on head- and footwear is the following. Instead of assuming that a fixed height can be added for all donors to account for head- and footwear, the addition may be randomized. That is, instead of Assumption 3 on head- and footwear one may assume that:

Assumption 4. For any person in the population, the height difference caused by head- and footwear has a $N(\nu,\eta^2)$ distribution.

As a consequence, the difference between the height of the perpetrator and suspect caused by differences in head- and footwear is normally distributed with mean 0 cm and variance $2\eta^2$ cm^2. Following the argumentation in the proof of Theorem 2, the formula obtained is the same as the one stated if the term $(1 + n^{-1})\ \sigma^2$ is changed into $(1 + n^{-1})\ \sigma^2 + 2\eta^2$ whenever it occurs.

5. Conclusions and discussion

There are several techniques for performing body height estimations on persons in digital images of crime scenes, cf. [1–6]. At the Netherlands Forensic Institute, the approach is by performing a reconstruction at the scene of crime. Using a 3D model of the crime scene and virtual replacement of the camera, operators perform height measurements by placing cylinders over the person in the image.

Variation between actual and measured heights of donors is decomposed into a systematic part (because of height loss by pose, 3D modeling of the scene of crime, operator biases) and a random part (due to natural variation). Using normal approximations and the observed variation on test persons, a method is described for obtaining confidence intervals for the height, including head- and footwear, of the perpetrator, described in Theorem 1 in Section 3. Since the number of test persons is usually limited, the result is in terms of the Student's t distribution.

In addition, an expression is obtained for the Likelihood Ratio, measuring the strength of evidence of resemblance of the actual height of a suspect and the measured height of the perpetrator. The result is stated in Theorem 2 and depends both on the rarity of the estimated perpetrator's height and on its closeness to the suspect's height.

The two technical theorems included in the Appendix describe how to estimate distributions in the case of limited sample sizes (numbers of test persons) and solve technicalities when calculating Likelihood Ratios. They may be relevant to similar problems in other forensic areas.

In [4], height estimations of persons in digital images are not limited to total body height but also performed on limbs and such. This results in various height measurements for each donor, of which it is suggested that it will have higher evidential value. However, the extra measurements may (1) be highly correlated and (2) have a much larger relative error margin than total body height. To investigate the latter, one would need large numbers of test persons.

The analysis of validation measurements described in the current paper does not depend on the method used and holds up as well if measurements are made on the basis of e.g. projective geometry (vanishing points).

Acknowledgements

The authors would like to thank Gerda Edelman, Bart Hoogeboom and Arnout Ruifrok for reviewing the paper.

Appendix A. Two statistical theorems

We turn to the proofs of the two technical theorems that were left in Sections 3 and 4.

A.1. Substitution of sample mean and variation for actual mean and variation when working with normally distributed variables

When working with normally distributed populations, we often encounter the problem that mean and variance of the normal distribution are unknown. Taking a sample from the population and estimating mean and variance of the population by those of the sample usually solve this problem. For small sample sizes, this introduces the difficulty that normalization by sample mean and variance do not exactly yield a standard normally distributed variable. We solve this using the following theorem.

Theorem A1. *Let $X_1, \ldots, X_n$ be a sample of independent variables sharing the same normal distribution, with unknown parameters. Let $\bar{X}$ and S_X^2 be the sample mean and sample variance. Then for any independent variable Y sharing the same distribution, the statistic*

$$V = \frac{Y - \bar{X}}{((1 + n^{-1})^{1/2} S_X)}$$

is Student's t-distributed with $n - 1$ degrees of freedom.

Proof. We denote mean and variance of the sample by μ and σ^2. Now $\bar{X} \sim N(\mu, \sigma^2/n)$, $Y \sim N(\mu, \sigma^2)$ and $\bar{X}$ and Y are independent. As a result, the variable

$$Z = \frac{Y - \bar{X}}{((1 + n^{-1})^{1/2} \sigma)}$$

is standard normally distributed.

Moreover, it is a standard result (see e.g. [9]) that $(n - 1)$ S_X^2/σ^2 has a χ^2 distribution with $n - 1$ degrees of freedom, and is independent of $\bar{X}$. Hence it is independent of Z as well, and the statement of the theorem follows from the definition of the Student's t distribution. □

A.2. Invariance of normal probability densities under certain linear transformations

The second technical theorem is sometimes useful when calculating probability densities in the LR definition, cf. [11].

Theorem A2. *Let the stochastic d-vector* **X** *have a multivariate normal distribution with mean* $\boldsymbol{\mu}$ *and covariance operator* **C**. *Let* **A** *be an invertible linear operator of the same dimensions as* **C** *such that* $|\det \mathbf{A}| = 1$. *Then with respect to the probability densities of the variables* **AX** *and* **X** *we have for all* **x** *that*

$$f_{\mathbf{AX}}(\mathbf{Ax}) = f_{\mathbf{X}}(\mathbf{x}).$$

Proof. Let **X** and **A** be as described. It is well-known (cf. [9], Chapter 8) that **AX** again has a (multivariate) normal distribution with mean $\mathbf{A}\boldsymbol{\mu}$ and covariance operator **ACA***, where by **A*** the transposed (adjoint) of **A** is meant. By the definition of the multivariate normal distribution then

$$f_{\mathbf{AX}}(\mathbf{Ax}) = (2\pi)^{-d/2}(det(\mathbf{ACA}^*))^{-1/2}\exp[-1/2\langle(\mathbf{ACA}^*)^{-1}(\mathbf{Ax} - \mathbf{A}\mu), \mathbf{Ax} - \mathbf{A}\mu\rangle],$$

with $\langle , \rangle$ denoting the Euclidian inner product, and a similar expression without **A** for $f_{\mathbf{X}}(\mathbf{x})$. Here, using standard computation rules for determinants:

$$det(\mathbf{ACA}^*) = det\mathbf{A}det\mathbf{C}det\mathbf{A}^* = (det\mathbf{A})^2 det\mathbf{C} = det\mathbf{C},$$

and since $\mathbf{A}$ is invertible:

$$(\mathbf{ACA}^*)^{-1}(\mathbf{Ax} - \mathbf{A}\mu) = (\mathbf{A}^*)^{-1}\mathbf{C}^{-1}\mathbf{A}^{-1}\mathbf{A}(\mathbf{x} - \mu)$$
$$= (\mathbf{A}^*)^{-1}\mathbf{C}^{-1}(\mathbf{x} - \mu).$$

Using the fact that $(\mathbf{A}^*)^{-1} = (\mathbf{A}^{-1})^*$ it then easily follows that $f_{\mathbf{AX}}(\mathbf{Ax}) = f_{\mathbf{X}}(\mathbf{x})$. □

Appendix B. Percentile points of the Student's *t* distribution

Table A1 gives percentile points of the Student's *t* distribution, for 2 up to 7 degrees of freedom. For a more extended table see e.g. [9].

Table A1
Percentile points for the Student's *t* distribution with *N* degrees of freedom, for $N = 2, \ldots, 7$

N	95%	97.5%	99%	99.5%
2	2.92	4.303	6.965	9.925
3	2.353	3.182	4.541	5.841
4	2.132	2.776	3.747	4.604
5	2.015	2.571	3.365	4.032
6	1.943	2.447	3.143	3.707
7	1.895	2.365	2.998	3.499

For example: if T_5 is *t*-distributed with 5 degrees of freedom, $P(T_5 \leq 3.365) = 0.99$.

References

[1] G. Edelman, I. Alberink, Estimation of body heights in digital images, Wiley Encyclopedia of Forensic Science, in press.

[2] D. Compton, C. Prance, M. Shears, C. Champod, A systematic approach to height interpretation from images, in: Proceedings of SPIE, vol. 4232, Technologies for Law Enforcement and Security, 2001, pp. 521–532.

[3] L. Klasen, O. Fahlander, Using videogrammetry and 3-D image reconstruction to identify crime suspects, in: Proceedings of SPIE, vol. 2942, Investigative Image Processing, 1996, pp. 162–169.

[4] N. Lynnerup, J. Vedel, Person identification by gait analysis and photogrammetry, J. Forensic Sci. 50 (2005) 112–118.

[5] A. Criminisi, A. Zisserman, L. van Gool, S. Bramble, D. Compton, A new approach to obtain height measurements from video, in: Proceedings of SPIE, vol. 3576, Investigation and Forensic Science Technologies, 1998, pp. 227–238.

[6] B. Hoogeboom, I. Alberink, M. Goos, Body height measurements in images. In preparation.

[7] F. Taroni, S. Bozza, A. Biedermann, Two items of evidence, no putative source: an inference problem in forensic intelligence, J. Forensic Sci. 51 (2006) 1350–1360.

[8] D. Lucy, Introduction to Statistics for Forensic Scientists, Wiley, Chichester, 2006.

[9] W. Pestman, Mathematical Statistics: An Introduction, De Gruyter, Berlin, Germany, 1999.

[10] C. Aitken, F. Taroni, Statistics and the Evaluation of Evidence for Forensic Scientists, 2nd ed., Wiley, Chichester, UK, 2004.

[11] D.V. Lindley, A problem in forensic science, Biometrika 64 (1977) 207–213.