

Estimating Production Time of Touchless Hand Drawing Gestures

Orlando Erazo^{1,2}, José A. Pino¹, Pedro Antunes³

¹ Department of Computer Science, University of Chile, Santiago, Chile
{oerazo, jpino}@dcc.uchile.cl

² Escuela de Informática, Universidad Técnica Estatal de Quevedo, Ecuador
oerazo@uteq.edu.ec

³ School of Information Management, Victoria University of Wellington, Wellington, New Zealand
pedro.antunes@vuw.ac.nz

Abstract. Quantitative user models such as CLC, Isokoski's and KLM have been used to estimate the production time of mouse and pen interactions (pointing, clicking, selecting, drawing, writing). In this paper, we assess if these models can be adapted to estimate the production time of touchless hand gestures (air figures of letters and numbers). New parameters were added to the existing models with empirical values drawn from experiments with users. Two metrics were used to evaluate model quality: strength of the relationship between estimated and observed times, and percentage root mean square error. The obtained results support the hypothesis that CLC, Isokoski's and KLM can be adapted to touchless hand gestures. The paper contributes with model modifications and parameters required to estimate the production times of touchless hand gestures.

Keywords: Touchless hand gestures; natural user interfaces; predictive evaluation; user models.

1 Introduction

Hardware devices such as Kinect, Leap Motion and Myo have contributed significantly to the emergence of a new class of graphical user interfaces labeled as Natural User Interfaces (NUIs), which most distinctive characteristic is using touchless hand gestures as a new interaction modality. Besides the obvious advantages in entertainment, NUIs are also becoming increasingly popular in fields such as education [1] and healthcare [2]. However, depending on each particular application, hand gestures may be very taxing on production time, i.e. the time spent by users engendering interactions with an application's user interface. This situation leaves user-interface designers with the difficult task of evaluating which touchless hand drawing gestures are most adequate and how they can be tailored to a user-interface under construction to optimize production time.

Considering the logistic difficulties of doing tests with real users, regarding planning, timing, laboratory setup, recruiting, and conducting experiments, a reasonable assumption is that user-interface designers may find value in instead adopting predictive evaluation. Predictive evaluation provides quantitative indications on how users may perform based on user models instead of real users, thus easing the abovementioned logistic problems.

Of course a consequence of adopting predictive evaluation is that we need quantitative user models. Although some quantitative models have already been proposed that include gestures [3, 4], they do not support the estimation of production time. However, an extensive body of research in quantitative user models for pointer- and pen-based gestures has already been developed. One of the most influential instances is the KLM (Keystroke Level Model) model originally proposed by Card, Moran and Newell and extended by others [5, 6, 7]. Another notable example is Fitts' Law [8], which has been widely used to estimate physical and virtual pointing [9, 10, 11, 12].

Some interactions like pointing, clicking and selecting can be done both with a mouse and in the air using hand gestures, and therefore can be modeled using KLM and Fitts' Law, as demonstrated by [9, 10, 11, 12]. However, more complex interactions such as drawing characters, numbers and figures require either extending the existing models with new parameters or developing whole new user models. Considering this landscape, the generic goal of the research reported in this paper is to extend the existing models to encompass hand gestures. More specifically, our goal is to extend the CLC (Curves, Line segments and Corners) [13], Isokoski's [14] and KLM [5] models to estimate the production times of touchless hand gestures. Our research is restricted to gestures performed by young adults in normal health conditions, with basic or no experience with touchless interactions, and using the dominant hand (fingers are not considered). The gestures of interest consist of drawing in the air figures or shapes such as letters or numbers.

2 Related Work

A commonly accepted strategy to evaluate the usability of a user interface is to use modeling techniques. This strategy presents the advantages of neither depending on real users nor requiring usability experts to participate in the evaluation process.

Models can vary in detail and complexity, ranging from *descriptive models*, which provide a framework for designers to delineate and reflect on usability problems, to *predictive models*, which use mathematical expressions for estimating user performance [15]. Unlike the descriptive category, predictive models can be used to objectively estimate the required time for performing a set of user interactions. Likewise the descriptive category, predictive models can be applied at early design stages, before starting to develop the real user interface.

Some notable user models have been proposed in the research literature. One of the most known and cited models is Fitts' Law [8], which allows estimating the time to point at a target based on the object size and distance. Taking into account that Fitts' Law is not adequate for certain types of tasks, and based on it, Accot and Zhai [16]

proposed the “steering law” for trajectory-based tasks. This model allows predicting the time to navigate through a two-dimensional tunnel, but it may not be adequate to analyze the trajectory of touchless hand gestures when there are no visual guides.

An alternative was proposed by Isokoski [14], who introduced a conceptually simple model that predicts production time for unistroke interactions done by expert users. According to this model, a gesture is first decomposed into a number of “needed straight-line segments” which are then counted to estimate the overall time-complexity of the gesture. The number of considered segments is the minimum necessary to make the gesture recognizable. Additionally, it is assumed that drawing a straight-line segment takes a constant time.

By comparing the model with real-user interactions, Isokoski [14] measured the *strength of the relationship between estimated and observed times* (usually noted as R^2) to be <0.85 . The *percentage root mean square error* (usually noted as %RMSE) of these measures was 30% [14].

Although the definition of “needed straight-line segments” is ambiguous [13, 17] (a procedure describing the reduction of gestures with curves into straight lines is missing), Isokoski’s model seems conceptually easy to extend to hand gestures because it requires estimating the constant time to produce a straight line gesture segment from experimental data.

Cao and Zhai [13] suggested a model to estimate the production time of single pen-stroke gestures. The model considers three features found in pen-stroke gestures: Curves, Line segments, and Corners (the reason why the model is referred to CLC).

For any gesture, the production time is calculated by summing up the estimated time durations of all gesture segments (see formula 1 below). The estimated production times of Curve and Line are defined in formulas 2 and 3. (Formula 4 can be used instead of formula 3.) The Corner, which is an abrupt change in stroke direction, was discarded by Cao and Zhai after empirical studies showed its insignificant impact on production times [13].

The CLC model reveals a strong relationship between estimated and observed times ($R^2 > 0.90$). Even though the model has been used in several research studies (e.g. [17, 18]), we do not have evidence that it has been used beyond pen-stroke gestures.

$$T = \sum T(line) + \sum T(corner) + \sum T(curve) \quad (1)$$

$$T(curve) = \frac{\alpha}{K} r^{1-\beta} \quad (2)$$

$$T(line) = mL^n \quad (3)$$

$$T(line) = aL + b \quad (4)$$

Where: α is the sweep angle; r is the radius of the arc; β and K are empirical constants. L is the length of the line; a , b , m and n are empirical constants.

Another well-known user model is the Keystroke Level Model (KLM) [5]. KLM defines a set of primitive operations: key press (K); point (P); button press (B); hand movement between keyboard and mouse (H); Drawing (D); and mental preparation (M). For each primitive operation, empirical studies with various types of users allowed to determine the respective average production time constant. The D primitive is the only one relevant to this study, even though it has several constraints worth noticing: drawing is done with the mouse, it only concerns straight-line segments, and it is assumed to be done on a square grid with 0.56cm. According to KLM, the production time of a drawing interaction is defined as a linear function of the number of segments (n_D) and the total length (l_D) of all segments (see formula 5) [5].

$$D(n_D, l_D) = an_D + bl_D \quad (5)$$

Where: a and b are constants ($a = 0.9$ and $b = 0.16$ in the original KLM version [5]).

All in all, the Isokoski's, CLC and KLM models advocate: 1) decompose a gesture either into a series of straight-line segments (Isokoski's and KLM) or into a series of curved and straight-line segments (CLC); 2) use a set of formulas and parameters to calculate the production time of each segment, adding them to obtain the overall production time of a gesture; and 3) derive the parameters from empirical studies with various users, so that the formulas adjust to reality. Nevertheless specific parameters for touchless hand gestures do not currently exist and will have to be further researched. We detail that process in the following section.

3 Hypotheses and Research Design

As mentioned above, there is no evidence that the three described models may be used for touchless hand gestures. Therefore, the following hypothesis should be tested:

H1: CLC, Isokoski's and KLM models can be adapted to predict the production time of hand gestures.

The starting point to test this hypothesis is the definition of formulas. Either the original formulas can be applied to hand gestures or they have to be extended to encompass the new conditions imposed by hand gestures. Hence, we introduce a first step in our study where the formulas are evaluated and adapted if necessary. This step is shown in Table 1 (step 1) and denoted as "An" (A - Adaptation).

After adapting the formulas, it is necessary to define new parameters for hand gestures. The second step requires carrying out several experiments with real users (En in table 1, step 2; E - Experiment) and then tuning the parameters so the models may reflect the users' performance.

At this point, the models should be ready to use. However, we still have to consider the quality of the estimations. Thus, we have to evaluate the models, that is, to verify the second hypothesis:

H2: The adapted models can predict the production time of hand gestures with acceptable quality.

We will test this hypothesis using the two metrics also adopted by Cao and Zhai [13], considering both the strength of the relationship between estimated and observed times (R^2), and percentage root mean square error (%RMSE). Furthermore, we will consider that a model has acceptable quality if R^2 and %RMSE are proximate to the values obtained by Cao and Zhai [13]: $R^2 > 0.90$ and %RMSE $< 30\%$.

H2 must be tested with a set of experiments with real users (E_n in Table 1, line 3). With a careful experimental setup, some of the experiments required to test H1 can also be used to test H2. That is the reason why we see in Table 1 that E1 and E2 are shared by steps 2 and 3. This is possible because: 1) the parameters required by CLC and Isokoski's models can be tuned using the same experiments; and 2) we use two different sets of users in E1 and E2, so that the users being used to tune a model can be reused to validate another model, but one set of users is not simultaneously used to tune and validate the same model.

In the remainder of this section we give more details about the experimental setup. The validation of H1 and H2 is discussed in detail in the following sections using the stepwise structure described in Table 1.

Table 1. Research design: adaptations of models (A) and experiments (E).

Steps	Models		
	CLC	Isokoski's	KLM
1. Definition of formulas	A1	A2	A3
2. Estimation of parameters	E1	E1	E2
3. Evaluation of models	E2	E2	E3

3.1 Apparatus and Method

The hardware setup for the experiments consisted of notebook, Kinect sensor and TV screen mounted in a controlled laboratory setting. The notebook was equipped with an i7 processor and 8GB of RAM. The Kinect sensor was used with a refresh rate of 30 fps, connected to track users' hand position and recognizing gestures, placed at a height of 0.9m and below the TV screen. The TV screen had 42in, 1360×768px resolution. The participants stood in an uncluttered space, 2.5m away from the Kinect sensor.

A custom software tool (Fig. 1) was developed for precisely controlling the experiments. The tool was developed using MS Visual C# and Kinect for Windows SDK V1.8 on Windows 7. The tool logged time marks and hand coordinates while a participant performs a gesture. The Dynamic Time Warping algorithm [19] was used for gesture recognition. The tool's interface consisted of an augmented video blending user interface controls and the real environment. Augmented video was adopted in order to avoid participants' distractions while performing the tasks. For instance, a person may judge his/her movements based on a hand cursor and try to make adjustments [20], especially because of the sensor noise [21] [10] [12] which should be

avoided. We fine-tuned these experimental conditions through a set of trial experiments.

The tool had two additional software modules focused on gesture analysis. The first module allowed recording gestures (Fig. 1-a), while the second one was able to reproduce every user-generated gesture using segmentation and logged hand coordinates (Fig. 1-b).

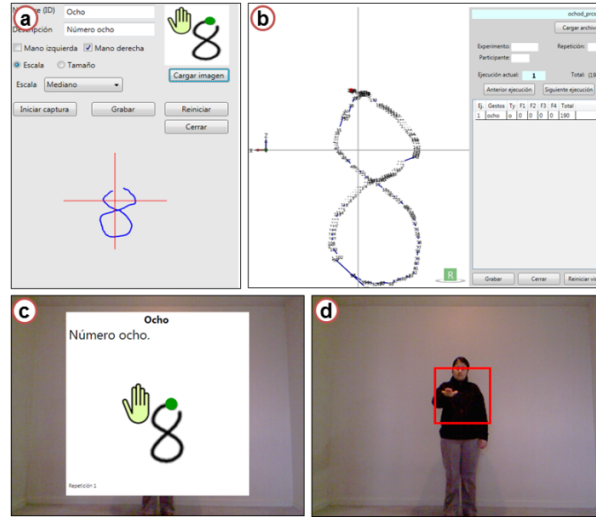


Fig. 1. Interface of the experimental software.

Each gesture instance was segmented in the phases proposed by [22] to measure the production times of gestures. More specifically, the measured stroke-phase time was defined as corresponding to production time, which does not account for the time spent by the participants in other gesture phases.

The participants in the experiments were University students (33 in total, aged between 17 and 28) invited by email, social networks, etc. The participants were not paid for their participation. Written informed consents were obtained before starting each experiment. A student was allowed to participate in a single experiment.

Before the experiment, each participant received written instructions and an explanation about the research goals. Then, the participant performed some training gestures guided by the software. When the participant indicated s/he was ready, more specific instructions appeared on the TV screen using a PowerPoint slideshow. Enough time was allowed to read the instructions. The instructions required that every gesture should be done inside a red box (gesture input area or gesture space, see Fig. 1-d), having approximately the same size, and balancing speed and accuracy. (The tool adjusted the size of the input area according to the size of the required gesture.) The instructions also noted the participants should use the dominant hand, and should start (preparation phase [22]) and finish (retraction phase [22]) a gesture with both hands in a relaxed position below hips.

The tool immediately started the data acquisition phase after displaying the instructions. The tool was programmed to randomly pick a gesture within a gesture set and display it for two seconds (Fig. 1-c). The gesture image was displayed along with a name and a very short description. After the two seconds period, the description disappeared, the red box was displayed, and the participant's gesture was collected (Fig. 1-d). When the gesture was correct, the tool displayed a green check and moved on to the next gesture. When a gesture was wrong, a red-cross mark was displayed, the input was discarded, and the participant had to re-enter the gesture.

Besides a practice session, every experiment included three blocks with gestures to be performed by the participants (e.g. a block only including straight lines). The specific characteristics of these blocks are defined in Sections 5 and 6. The tool included a resting period between blocks of gestures.

4 Definition of Formulas

This section discusses the definition and adaptation of each model to hand gestures. The tuning and validation of formulas are explained later.

4.1 CLC Model

Although Cao and Zhai [13] adopted several formulas (1, 2, 3 and 4) for the CLC model, there are other options that might be used to improve predictions. Based on regression analysis (discussed later), we suggest that formula 6 can be used instead of formula 2 to estimate the production time of curves. Another simpler formula that may be applied to curves, which Cao and Zhai did not test, is a linear function of the curve's radius and angle (formula 7). These two formulas may contribute to reduce the %RMSE of CLC, but they should be tested against the original one.

$$T(\text{curve}) = \frac{\alpha^a}{K} r^{1-\beta} \quad (6)$$

$$T(\text{curve}) = m r + n \alpha \quad (7)$$

Where: α is the sweep angle; r is the radius of the curve; a , β , K , m and n are empirical constants.

4.2 Isokoski's Model

Given the conceptual simplicity of Isokoski's model [14], we consider it may be straightforwardly adapted to hand gestures. Formula 8 may be applied bearing in mind that hand gestures will be reduced into a sequence of straight lines and an empirical constant is necessary to estimate the time taken to generate every straight line. This reasoning has two implications. First, Isokoski did not provide a constant time for performing a straight-line segment, so we have to estimate that constant. Second,

Isokoski' does not provide a clear procedure to reduce curves into straight lines, which may range between 1 (too much error) and an arbitrary large number (less error but more difficult to calculate). We adopt the procedure suggested by Vatavu et al.: "if the angle α inscribed by an arc was greater than 270° use 3 segments; if $\alpha < 120^\circ$ use 1 segment; otherwise use 2 segments" [17] (p. 97).

$$T = \text{\#segments} * \text{constant_time} \quad (8)$$

4.3 KLM

As already mentioned, we are only using the D operator of KLM, and hence using formula 5 to estimate production time of hand gestures. Analyzing the original definition in detail [5], we note that formula 5 may not be applicable to gestures with curves and corners. We suggest that curves be approximated to straight-line segments by applying the procedure adopted by Vatavu et al. [17] (p. 97). Regarding corners, we suggest counting the number of corners (n_C) multiplied by an empirical constant, as shown in formula 9.

$$D_c(n_D, l_D, n_C) = a n_D + b l_D + c n_C \quad (9)$$

Where a , b and c are empirical constants.

5 Estimation of Parameters

In this section we give further insights about formulas 6 to 9 delineated above. The discussion is organized in two steps. In the first step, we describe the experiments (E1 and E2, see Table 1) conducted to obtain empirical data about user's hand gestures. In the second step we present the final formulas with the estimated empirical constants.

5.1 CLC Model

Experiment. We repeated most of the experimental process described by Cao and Zhai [13] using hand gestures to obtain empirical constants for formulas 2, 3, 4, 6 and 7. Experiment E1 involved gathering data for three gesture components: straight lines, curves and corners. This experiment then implied configuring the tool discussed in Section 3 to request participants to produce variations of these individual gesture components (see below). Each participant would produce the same gesture three times in order to increase precision. Trying to avoid learning and/or sequence effects, the order of components was counterbalanced.

Various lengths ($L = \{0.4, 0.6, 0.8\}$ meters in motor space) and orientations (0, 45, 90 and 135 degrees counterclockwise) were tested when producing straight lines. For curves, various radiuses ($r = \{0.2, 0.3, 0.4\}$ meters in motor space) and sweep angles ($\alpha = \{90, 180, 360\}$ degrees) were tested. Start angle (90 degrees) and direction (clockwise) were treated as control variables for curve gestures. Various corner angles

($\theta = \{45, 90, 135\}$ degrees) and directions (CW and CCW) were tested to produce corners. Length was kept constant (0.6cm in motor space). Twelve persons participated in E1 (mean age 21y, $\sigma = 2$).

Results. In general, the results obtained from E1 for the CLC model have similar significance to those obtained by Cao and Zhai [13], which provides a first indication that CLC can be used for estimating production time of touchless hand gestures. In detail:

Straight Lines. We observed statistically significant differences when measuring production time and varying length ($F_{2,22} = 10.47$, $p < 0.05$) and orientation ($F_{3,33} = 2.92$, $p < 0.05$). No significant length \times orientation interaction effects were found ($F_{6,66} = 0.32$, ns). Fig. 2 shows the relation between length and production time for each orientation. We note that Cao and Zhai [13] did not take into account orientation in their estimations because its effect was considered smaller than length. We made the same decision due to the similarity of our results, but we also computed the correlation coefficients to confirm it. We found no correlation between orientation and time production ($r = -0.022$). Finally, performing regression analysis of our experimental data, we obtained the empirical constants shown in formulas 10 and 11.

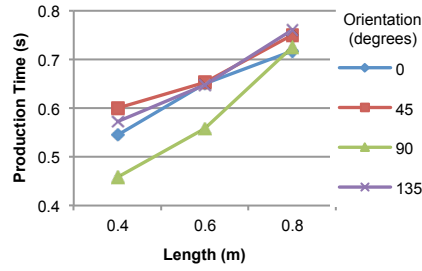


Fig. 2. Straight line production time.

$$T(\text{line}) = 0.486 L + 0.345 \quad (R^2 = 0.796) \quad (10)$$

$$T(\text{line}) = 0.803 L^{0.442} \quad (R^2 = 0.746) \quad (11)$$

Where L and T are given in meters and seconds respectively.

Curves. As we expected, measured differences in production time were statistically significant for both radius ($F_{2,22} = 12.33$, $p < 0.05$) and angle ($F_{2,22} = 110$, $p < 0.05$). We found significant radius \times angle interaction effects ($F_{4,44} = 3.72$, $p < 0.05$). Fig. 3 shows the relation between sweep angles and production time for each orientation radius. After performing regression analysis with our experimental data, we obtained formulas 12, 13 and 14 below.

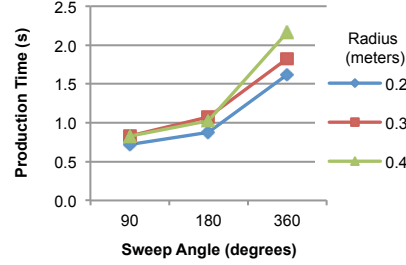


Fig. 3. Curve production time.

$$T(\text{curve}) = \frac{\alpha}{1.939} r^{1-0.711} \quad (12)$$

$$T(\text{curve}) = \frac{\alpha^{0.615}}{1.249} r^{1-0.711} \quad (R^2 = 0.919) \quad (13)$$

$$T(\text{curve}) = 1.338 r + 0.236 \alpha \quad (R^2 = 0.942) \quad (14)$$

Where α , r and T are given in radians, meters and seconds respectively.

Corners. Following Cao and Zhai's method, we computed the “net contribution time” of corners [13]. Thus, for our experimental data and using the average time to perform a line 0.6cm long (see Table 2): $T(\text{corner}) = \text{sample production time} - 2 * 0.627$ (seconds).

The measured differences in production time were statistically significant for corner angle ($F_{2,22} = 6.49$, $p < 0.05$), but not for direction ($F_{1,11} = 2.05$, $p > 0.05$). We found no significant angle \times direction interaction effects ($F_{2,22} = 0.24$, ns). Taking into account that the average $T(\text{corner})$ seems to fluctuate around zero (Fig. 4), we made a deliberate simplification (formula 15): to omit corners in the model (Cao and Zhai [13] made the same decision). Although these results confirm previous preliminary findings, which postulate that corners have influence on production time of hand gestures [4], we think further research is necessary to adequately model the impact of corners in hand gestures.

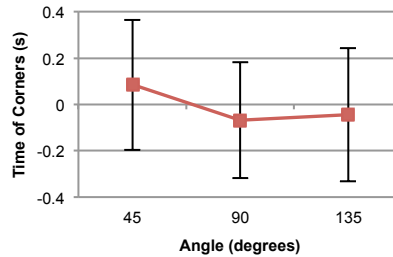


Fig. 4. Net corner time contribution. Error bars indicate 1 SD.

$$T = \sum T(line) + \sum T(curve) \quad (15)$$

5.2 Isokoski's Model

Experiment. As mentioned above, data obtained from experiment E1 was reused to build an estimation model based on Isokoski's proposal [14]. Although Cao and Zhai [13] state that a constant time model should be invalidated, we nevertheless decided to build this model because of its conceptual simplicity. An average time was calculated for each straight line produced by the participants in the experiment (Table 2). Moreover, we estimated a fourth value to evaluate the model with a smaller straight-line segment (0.2m). These times must be verified for selecting the best one by using different gestures (see next section).

Results.

Table 2. Constant times for Isokoski's model.

Straight line lengths, $L(m)$	Observed time, $t(s)$	SD(s)
0.2	0.442	NA
0.4	0.544	0.111
0.6	0.627	0.180
0.8	0.738	0.219

5.3 KLM

Experiment. Taking into account that experiment E1 is focused on curves, straight lines and corners, we had to perform another experiment (E2) to estimate the empirical constant for the D operator of KLM, given that it is based on the number of segments and the total length of all segments. The experiment consisted of drawing 14 gestures¹ (Fig. 5) in random order. Gestures were performed inside the gesture space, which was a 0.6m square. Twelve persons took part in E2 (mean age 23y, $\sigma = 2$).

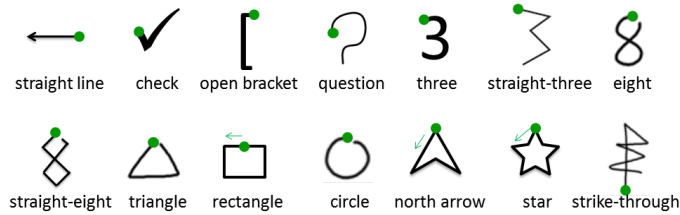


Fig. 5. Gestures used in experiment 2.

¹ The green point indicates the start of the gesture. Strike-through was selected from [17].

Results. The procedure adopted by Card et al. [5] for drawing straight-line segments (formula 5) was tested with hand gestures. This means that gestures with curves (“question”, “three”, “eight” and “circle”, in Fig. 5) were not used to build the model. The number of segments (n_D) of each gesture produce by participants in the experiment was counted and the total length (l_D) of each gesture was computed (geometrically). Formula 16 was obtained by performing regression analysis. The resulting R^2 value was high (0.988), but we thought the model could still be improved. We obtained a higher R^2 value (0.99, see formula 17) considering the number of corners and using formula 9. (Corners were counted depending on the gesture start point.)

$$D(n_D, l_D) = 0.386 n_D + 0.349 l_D \quad (R^2 = 0.988) \quad (16)$$

$$D_c(n_D, l_D, n_c) = 0.223 n_D + 0.297 l_D + 0.173 n_c \quad (R^2 = 0.99) \quad (17)$$

Formula 16 was then tested to estimate the production times of the four gestures with curves shown in Fig 5, which had not been previously used to estimate the empirical constants (condition D in Fig. 6). Additionally, formula 17 was tested with and without corners (conditions D_c and D_c^*). The obtained results for these three conditions are shown in Fig. 6. These results indicate that these models can also be applied to gestures with curves.

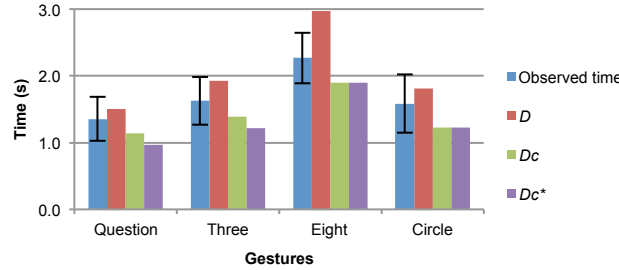


Fig. 6. Comparison of observed and predicted times of 4 gestures with curves using the D , D_c and D_c^* conditions. Error bars indicate 1 SD.

6 Evaluation of Models

The production times of real hand gestures must be compared against predicted values in order to evaluate the adapted models. We tried to reduce the number of experiments to a minimum and thus, we decided to reuse data from experiment E2 to evaluate the adapted CLC and Isokoski’s models, whose parameters were developed from E1 using a different cohort and different gestures. Regarding the evaluation of the adapted KLM model, a new experiment had to be setup (E3), since E2 was used to estimate the parameters for this model.

6.1 CLC Model

Formula 15 was suggested to estimate production time using the CLC model, with the provision that formulas (10, 11) and (12, 13, 14) can be considered options for measuring straight-lines and curves, respectively. Six possibilities can be analyzed to identify the best estimation approach by combining these formulas.

The results obtained from E2 are shown in Table 3 for the six formula combinations. We note that some R^2 values are lower than the baseline (Cao and Zhai's results [13]), but the obtained %RMSE values are better. Furthermore, the differences between estimates are relatively small. The best results are obtained using the linear model for straight lines (formula 10) and the modified model for curves (formula 13). Fig. 7 displays the predicted versus observed data using this formula combination.

Table 3. Comparison of CLC model predictions.

Formulas	R^2	%RMSE
(15), (10) and (12)	0.834	18.8
(15), (11) and (12)	0.779	18.8
(15), (10) and (13)	0.859	15.7
(15), (11) and (13)	0.810	15.7
(15), (10) and (14)	0.849	16.9
(15), (11) and (14)	0.798	16.9

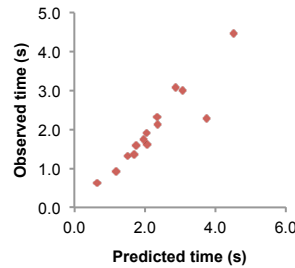


Fig. 7. CLC model prediction.

6.2 Isokoski's Model

E2 also allowed validating the Isokoski's model expressed in formula 8 with the empirical constants defined in section 5.2. The obtained results, shown in Fig. 8, suggest that selecting a constant straight-line length of 0.4m gives the least estimation error. Fig. 9 shows the relationship between predicted and measured production times for the suggested straight-line length ($R^2 = 0.935$, $L = 0.4\text{m}$, $t = 0.544\text{s}$).

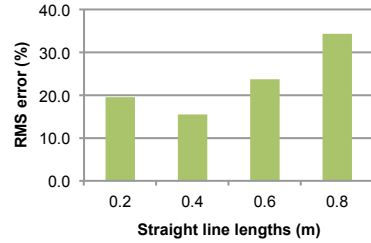


Fig. 8. Comparison of Isokoski's model prediction errors.

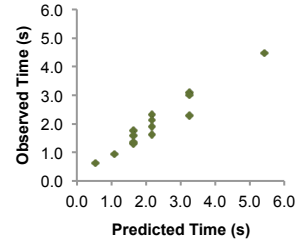


Fig. 9. Isokoski's model prediction.

Finally, we compared the measured production time with the best results estimated by the CLC and Isokoski's models (Fig. 10). Isokoski's model is slightly better than CLC, but the difference is quite small to choose the best one. Also, we note the worst predictions were made for gestures "three" and "eight", which are outside ± 1 SD.

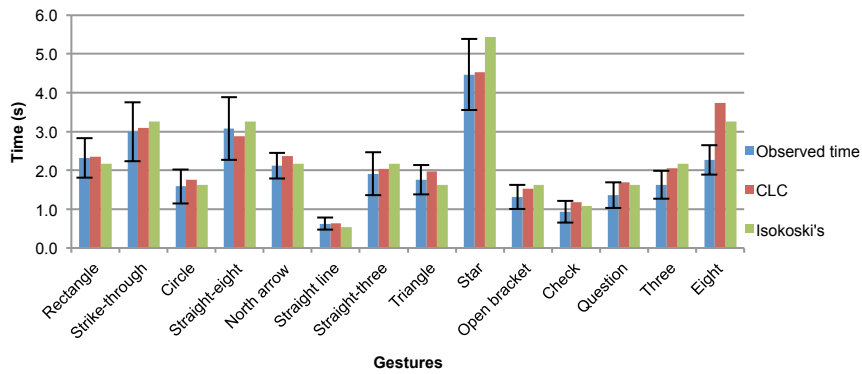


Fig. 10. Comparison of observed and predicted times using both CLC and Isokoski's models. Error bars indicate ± 1 SD.

6.3 KLM

Experiment E3 was set up in a similar way to E2. Nine participants (mean age 21y, $\sigma = 3$), performed the 6 gestures² shown in Fig. 11.



Fig. 11. Gestures used in E3.

² The green point indicates the start of the gesture. Steep-hill was selected from [17].

Stroke times were compared against predicted values using formulas 16 and 17. Before applying the formulas, gestures with curves (“5”, “E” and “steep-hill”) were reduced into straight lines. Additionally, formula 17 was calculated with and without corners (D_c and D_c^* conditions). For instance, gesture “E” was evaluated using 1 and 0 corners. The obtained results are shown in Fig. 12.

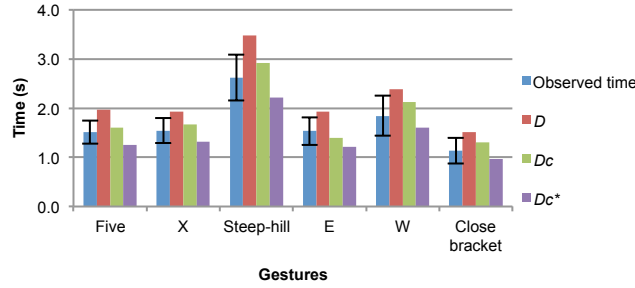


Fig. 12. Comparison of observed and predicted times for experiment E3 using D , D_c and D_c^* conditions. Error bars indicate 1 SD.

The highest R^2 (0.995) value was observed for the D condition, while the lowest (0.947) was observed for D_c , even though they were quite approximate. Since the %RMSE was favorable to the D_c condition (10.4), we suggest that D_c could be considered the best one overall.

7 General Comparison

In this section we finally compare the three models, using again the data collected in experiment E3 and focused on the formulas and empirical constants that produced the best estimates.

Table 4 shows the selected formulas and quality of estimates using the two quality criteria adopted by this study. Fig. 13 provides a more detailed comparison using the observed and predicted production times for the six gestures used to evaluate the estimation models. The highest R^2 value was obtained for the CLC model, but the differences to KLM (D_c condition) are quite small. On the other hand, the lowest %RMSE was obtained with KLM (D_c). Consequently, we suggest using KLM (D_c) to predict the production time of hand gestures.

Table 4. Comparison of the three models.

Model name	Formulas	R^2	%RMSE
CLC	(15), (10) and (13)	0.996	25.7
Isokoski's	#segments * 0.544	0.881	25.7
KLM (D_c)	(17)	0.947	10.4

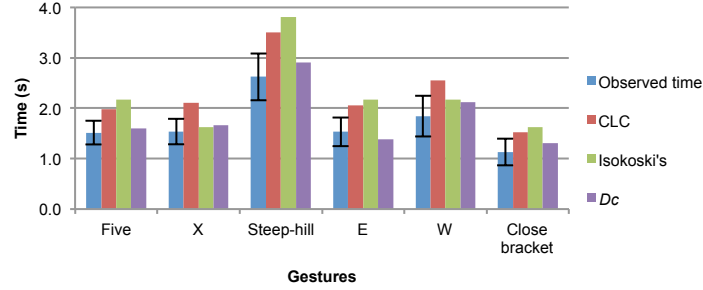


Fig. 13. General comparison of observed and predicted times using CLC, Isokoski's and KLM models. Error bars indicate 1 SD.

8 Discussion and Conclusions

In this paper we analyze three estimation models for predicting the production time of users' interactions with other types of user interfaces. We extend these models to hand gestures. Empirical experiments were accomplished to tune and validate the models. The quality of the estimates was evaluated using two criteria: strength of the relationship between estimated and observed times, and percentage root mean square error. In a broad perspective, we can conclude that the three models can be used with hand gestures, which confirms hypothesis H1. Furthermore, we provide new or updated formulas and empirical constants required to use the models with hand gestures.

The constant-time estimation model, which was proposed by Isokoski for unistroke writing [14], is the simplest model. This model is very easy to use because it reduces gestures to straight-line segments, counts them, and uses a constant multiplier that reflects the average time necessary to produce a straight-line segment. The constant multiplier depends on the constant length assigned to a segment.

According to our results, if gestures are drawn inside a square gesture space with 0.6m sides, acceptable results can be achieved with a segment that is 0.4m long. Conversely, the simplifications required by this model lead to erroneous estimations when using variable gesture spaces (i.e. making gestures of different sizes).

An alternative approach, which we also analyzed, consists in using the CLC model, which breaks down gestures into curves, lines and corners [13]. This model avoids reducing curves and corners to straight-line segments.

We provide new or updated formulas and empirical constants required to use the CLC model with hand gestures. Moreover, our experiments indicate that corners influence production time and therefore should not be neglected. Additionally, slightly different formulas were evaluated, leading us to suggest a new formula for estimating the production time of hand gestures using CLC.

The KLM model also reveals easy to adapt to hand gestures, because it is only based on the number of straight-line segments and the total length of a gesture. Conversely, this strategy also reveals a limitation because KLM's D operator does not

take into account other components like corners and curves. Trying to overcome these limitations, we included corners as a third parameter in KLM's estimation formula. The experimental results indicate these modifications provide good results.

Gestures with curves were analyzed as if they were straight lines and with the options of counting or not the number of corners. The obtained results show that counting corners improves the quality of the estimation. Consequently, the adapted KLM formula we suggest counts the number of segments, the total length and the number of corners of a gesture.

Regarding the experiments, we should note the following. First, we could observe a relatively high variation of gesture production times among participants. Although we did not compute a global or final value, the coefficient of variation is, on the average, about 30%. Second, the models were adapted and evaluated only using the gestures' stroke phase, even though they could also be analyzed taking into account a more comprehensive view (e.g. [22]). Third, the gestures used in our experiments were performed using only the dominant hand, although users may perform gestures with the other hand [23]. This constraint may have an effect on the estimates (e.g. [12]).

The model we suggest as the best to estimate production time of hand gestures obtained $R^2 \geq 0.947$ and $\%RMSE = 10.4$, which are better than the ones obtained by Cao and Zhai [13] for single pen-stroke gestures. Regarding hypothesis H2, we observe it is validated for the CLC and KLM (D_c condition).

We expect to conduct more evaluations in the future with more users and more gestures. We also consider studying other hand gesture types like hover, tap and swipe.

Acknowledgments

Orlando Erazo appreciates the financial support provided for this research by SENESCYT, Ecuador.

References

1. Blum, T., Kleeberger, V., Bichlmeier, C., Navab, N.: miracle: An Augmented Reality Magic Mirror System for Anatomy Education. In: 2012 IEEE Virtual Reality Short Papers and Posters, pp. 115–116. IEEE Press (2012)
2. Erazo, O., Pino, J. A., Pino, R., Fernández, C.: Magic Mirror for Neurorehabilitation of People with Upper Limb Dysfunction Using Kinect. In: 47th Hawaii International Conference on System Sciences, pp. 2607–2615. IEEE Press (2014)
3. Barclay, K., Wei, D., Lutteroth, C., Sheehan, R.: A Quantitative Quality Model for Gesture Based User Interfaces. In: Proceedings of OzCHI 2011, pp. 31–39. ACM Press, New York (2011)
4. Erazo, O., Pino, J. A.: Estimating the Difficulty of Touchless Hand Gestures. IEEE Latin America Transactions 12, 17–22 (2014)
5. Card, S., Moran, T., A. Newell: The Keystroke-Level Model for User Performance Time with Interactive Systems. Communications of the ACM 23, 396–410 (1980)

6. Holleis, P., Otto, F., Hussmann, H., Schmidt, A.: Keystroke-Level Model for Advanced Mobile Phone Interaction. In: Proceedings of CHI 2007, pp. 1505–1514. ACM Press, New York (2007)
7. Erazo, O., Pino, J. A.: Predicting Task Execution Time on Natural User Interfaces based on Touchless Hand Gestures. In: Proceedings of IUI 2015, pp. 97–109. ACM Press, New York (2015)
8. Fitts, P. M.: The information capacity of the human motor system in controlling the amplitude of movement. *J. of Exp. Psychology* 47, 381–391 (1954)
9. Pino, A., Tzemis, E., Ioannou, N., Kouroupetroglou, G.: Using Kinect for 2D and 3D Pointing Tasks: Performance Evaluation. In: Kurosu, M. (ed.) HCII 2013. LNCS vol. 8007, pp. 358–367. Springer, Heidelberg (2013)
10. Sambrooks, L., Wilkinson, B.: Comparison of gestural, touch, and mouse interaction with Fitts' law. In: Proceedings of OzCHI 2013, pp. 119–122. ACM Press, New York (2011)
11. Schwaller, M., Lalanne, D.: Pointing in the Air: Measuring the Effect of Hand Selection Strategies on Performance and Effort. In: Holzinger, A., Ziefle, M., Hitz, M., Debevc, M. (eds.) SouthCHI 2013. LNCS vol. 7946, pp. 732–747. Springer, Heidelberg (2013)
12. Zeng, X., Hedge, A., Guimbretiere, F.: Fitts' Law in 3D Space with Coordinated Hand Movements. In: Human Factors and Ergonomics Society Annual Meeting. SAGE Publications (2012)
13. Cao, X., Zhai, S.: Modeling Human Performance of Pen Stroke Gestures. In: Proceedings of CHI 2007, pp. 1495–1504. ACM Press, New York (2007)
14. Isokoski, P.: Model for Unistroke Writing Time. In: Proceedings of CHI 2001, pp. 357–364. ACM Press, New York (2001)
15. MacKenzie, I.S.: Motor Behavior Models for Human-Computer Interaction. In: HCI Models, Theories, and Frameworks, pp. 27–54. Morgan Kaufmann, San Francisco (2003)
16. Accot, J., Zhai, S.: Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks. In: Proceedings of CHI 1997. ACM Press, New York (2001)
17. Vatavu, R., Vogel, D., Casiez, G., Grisoni, L.: Estimating the Perceived Difficulty of Pen Gestures. In: Campos, P., Graham, N., Jorge, J., Nunes, N., Palanque, P., Winckler, M. (eds.) Interact 2011. LNCS vol. 6947, pp. 89–106. Springer, Heidelberg (2011)
18. Tu, H., Ren, X., Zhai, S.: A comparative evaluation of finger and pen stroke gestures. In: Proceedings of CHI 2012, pp. 1287–1296. ACM Press, New York (2012)
19. Senin, P.: Dynamic time warping algorithm review. Information and Computer Science Department, University of Hawaii at Manoa Honolulu (2008)
20. Sutter, C., Müsseler, J., Bardos, L., Ballagas, R., Borchers, J.: The impact of gain change on perceiving one's own actions. *Mensch & Computer*, 147–156 (2008)
21. Livingston, M. A., Sebastian, J., Ai, Z., Decker, J. W.: Performance Measurements for the Microsoft Kinect Skeleton. In: 2012 IEEE Virtual Reality Short Papers and Posters, pp. 119–120. IEEE Press (2012)
22. McNeill, D.: Guide to Gesture Classification, Transcription, and Distribution. In: *Hand and Mind: What Gestures Reveal about Thought*, pp. 75–104. The University of Chicago Press (1992)
23. Annett, M., Bischof, W.: Your Left Hand Can Do It Too! Investigating Intermanual, Symmetric Gesture Transfer on Touchscreens. In: Proceedings of CHI 2013, pp. 1119–1128. ACM Press, New York (2013)