Geometry and Gesture-based Features from Saccadic Eye-Movement as a Biometric in Radiology

Folami T. Alamudun*, Tracy Hammond, Hongjun Yoon, and Georgia D. Tourassi

Biomedical Sciences, Engineering, & Computing Group,
Oak Ridge National Laboratory, Oak Ridge, TN, USA 37831
Sketch Recognition Lab, Department of Computer Science & Engr.,
Texas A&M University, College Station, TX, USA, 77843
alamudunft@ornl.gov,hammond@cs.tamu.edu,
{yoonh,tourassig}@ornl.gov
http://www.ornl.gov/bsec,http://srl.tamu.edu

Abstract. In this study, we present a novel application of sketch gesture recognition on eye-movement for biometric identification and estimating task expertise. The study was performed for the task of mammographic screening with simultaneous viewing of four coordinated breast views as typically done in clinical practice. Eye-tracking data and diagnostic decisions collected for 100 mammographic cases (25 normal, 25 benign, 50 malignant) and 10 readers (three board certified radiologists and seven radiology residents), formed the corpus for this study. Sketch gesture recognition techniques were employed to extract geometric and gesture-based features from saccadic eye-movements. Our results show that saccadic eye-movement, characterized using sketch-based features, result in more accurate models for predicting individual identity and level of expertise than more traditional eye-tracking features.

Keywords: eye-tracking, biometrics, sketch recognition, mammography

1 Introduction

Survival of *breast cancer* disease is largely dependent on early detection through the annually recommended mammographic screening process. Studies show that

* This manuscript has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. De- partment of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (http://energy.gov/downloads/doe-public-access-plan).

through early detection, while the disease is localized, patients have a 98.5% relative survival rate in comparison to a 25% survival rate when the cancer is metastasized; a point at which the disease becomes incurable [42].

The timely detection of breast cancer is made possible through a process known as mammographic cancer screening. Mammographic screening is a specialized examination of X-ray images of interior breast tissues by a trained radiologist. Achieving expertise in radiology requires specialized training, which consists of 5 - 7 years of Radiology residency and fellowship, and years of experience during which the practitioner develops an *intuition* for the task. Expert radiologists exhibit notably outstanding characteristics, such as increased speed and higher overall accuracy with which he/she makes decisions on the pathology of an image, which differentiate them from non-experts. However, the length of training, and the specific nature and duration experience necessary to achieve expertise has been the subject of much research in medical imaging [5,30,23].

Although the exact relationship between experience and expertise remains unclear, one approach to establishing a quantitative relationship between the two, within the context of mammography, is through identifying differences in visual search behavior between experts and non-expert image readers [30,23]. In a study of six image readers (board certified radiologists and Radiology residents), Krupinski [22] compared cumulative cluster dwell times on 20 mammographic cases between experience groups. A comparison of the median values for experienced and inexperienced image readers revealed that experienced readers tend to have shorter dwell times. Their findings suggest that temporal measures of visual search behavior may be important factors in differentiating experience level of image readers.

Kundel and LaFolette [25] evaluated the eye-movements of 24 subjects, which included laymen, medical students, and experienced radiologists while viewing normal and abnormal chest radiographs. They reported an evolution of observers' scanpaths from the localized central patterns of first-year medical students to the circumferential patterns of the experienced radiologist. They noted that, in addition to the distinct nature of experienced radiologists' scanning patterns, experienced radiologists also moved their eyes to the target faster and were more accurate at interpreting what they saw. Kundel and LaFolette's findings suggest that *geometric properties* of scanning patterns formed during visual search may be important factors in differentiating between experienced and inexperienced image readers.

To investigate human factors associated with proficiency of diagnostic pathology, Krupinski et al. [24] conducted a study examining the eye-movement of nine image readers of varied experience level (medical students, Pathology residents, and pathologists). They reported that, when compared with Radiology residents and medical students, experienced pathologists exhibited longer saccades on average (measured in seconds). A similar trend was noted when comparing medical students Radiology residents.

In addition, they reported that the average saccade velocity for experienced pathologists was lower in comparison with Radiology residents, who's average velocity was higher than those recorded for medical students. They noted that the decreasing trend in saccade velocity with years of experience was consistent within the experienced pathology group (board certified pathologists), with the more experienced pathologists exhibiting a significantly lower average saccade velocity than the less experienced pathologists. Krupinski et al's findings suggest that distance and velocity measures of eye-movement during visual search in diagnostic pathology may also be important factors in differentiating between experienced and inexperienced readers.

In this paper, we describe a novel application of sketch gesture recognition to extract discriminative information from eye-tracking data for the purpose of user identification and for determination of task proficiency in Radiology. The remainder of this paper is organized as follows. Section 2 provides a general introduction to the domain of sketch recognition along with related. Section 2 also covers related work in eye movement-based biometric identification. Section 3 describes our experimental procedure and data processing methods. Section 4 presents the results from our experiments. Sections 5 gives a brief discussion of results followed by conclusions and acknowledgements in Sections 6 and 7 respectively.

2 Related Work

2.1 Sketch Gesture Recognition

Sketch is considered a natural form of communication involving free form shapes, letters, and numbers, which encode contextual meaning. Sketches can be considered as a special class of gestures. The fundament in sketch recognition involves encoding patterns contained within a sketch gesture in a manner, which permits accurate interpretation and inference based on the intent of the author of the sketch gesture [16]. The domain of sketch recognition utilizes machine intelligence to capture and interpret intent of the author making the sketch gestures. The correct interpretation of gesture intent enables the integration of sketch gestures in user interface systems, which in turn enables intelligent manipulation and computation on the recognized input.

There are numerous algorithmic contributions to general artificial intelligence from the domain of sketch recognition. The majority of sketch recognition algorithms fall into one of three broad categories: geometry-based algorithms [34], vision-based (appearance-based) recognition algorithms [20,32], and gesture-based (motion-based) algorithms [40,27], or hybrid combinations of these [7].

Geometry-based algorithms apply geometric relationships and constraints to describe primitive (basic) shapes, which combine to form recognizable high-level shapes [34]. Appearance-based recognition algorithms rely on the appearance of a sketched shape; ignoring timing and ordering constraints of data points [38]. These algorithms rely on recognition techniques, such as template-matching, on the snapshot of a sketched shape to distinguish between shapes [20,32]. Gesture-based (motion-based) recognition algorithms rely primarily on the path of motion of a strokes that make up a sketch shape. Gesture-based algorithms characterize shapes based on how individual strokes are drawn (the path of each stroke) in contrast with the shape of the stroke, even though the latter can be correlated. These types of algorithms were initially conceptualized for identification of a small set of application-specific gesture commands [40,27]. Rubine [40] developed a pen input gesture-based recognition system (GRANDMA), which enabled recognition of single stroke gestures through simple trainable linear classifiers. In this work, Rubine proposed and evaluated 13 features for classifying ten different gesture datasets, each containing 15 classes, and reported an average accuracy of 98%. In a followup work, Long et al. [27] proposed 11 additional features to those developed by Rubine.

Sketch recognition algorithms were previously applied to solve challenging pattern recognition problems in other domains [14,43,26]. Dixon and Hammond in 2010 [9,15] and Pramanik and Bhattacharjee in 2012 [36] applied sketch recognition algorithms to identify faces in images from sketched drawings. They reported an average of 86% similarity with the top five matches using their method, which was significantly higher than averages from the two alternatives presented (eigenface: 43%, and sketch transform method: 80%).

Cig and Sezgin [6] developed a eye-movement interaction system, which interprets eye-movement patterns as auxiliary commands to augment pen-based gestures as a mode selection mechanism (drag, minimize, scroll etc.) during sketch interaction. Their results demonstrated that manipulation commands can be recognized with 88% accuracy using natural gaze behavior during pen interactions. In [32], Ouyang and Davis presented a robust, multiple domain sketch recognition system, which uses vision based decomposition methods to classify hand-drawn symbols. Their system represented symbols as a set of feature images, in contrast to geometric or temporally ordered data points. These image features capture properties of the constituent strokes in a sketch symbol, such as orientation and the location of end points.

More advanced systems [13,45] are able to identify high-level shapes by using geometry-based algorithms to characterize its constituent low-level shapes. Valentine et al. developed *Mechanix*, an intelligent, interactive, on-line tutoring system, which allows engineering students to enter planar truss and free-body diagram solutions to homework problems [45]. The work reported in this paper does not represent the first time sketch-based features have been applied to human motions other than pen [33,29,2], but it is the first time they have been applied to characterize eye-movement.

2.2 Eye-Movement as a Biometric

Biometrics refer to authentication techniques, which rely on easily verifiable physical characteristics of an individual. Biometric *identifiers* are categorized as measurable physiological and behavioral properties of the individual. Physiological characteristics are measures related to some property of the physical body, which include fingerprint, footprint, palmprint, palm veins, face, DNA, iris, and retina. Behavioral characteristics are measures specific the behavior of a person (behaviometrics), which include typing cadence, gait, hand-writing, and voice. Eye-movements do not easily lend themselves to forgery, since they are largely dependent on brain activity and extra-ocular muscle characteristics, which are *unique* to the individual not unlike the biomechanics of walking (gait). This property makes eye-movement an attractive option for biometric identification.

In a previous work, Noton and Stark [31] observed that individuals tend to repeat certain scanpath trajectories during repeated viewings of a given pattern. In their experiments, they tested this theory, coined *scanpath theory*, and found that the general scanpath for a subject during a first viewing of a pattern was repeated in initial eye-movements of approximately two-thirds (65%) of subsequent viewings. In addition, Noton and Stark observed that the scanpath produced by an individual for a given stimulus pattern was unique and varied for each subject [31]. These findings were also supported by subsequent research in reading related information processing [39,41].

Eye-movements were first explored as a potential biometric identifier in [21]. In this work, Kasprowski and Ober used a combination of eye reaction time (the period of time between introduction of stimulus and eye reaction), and stabilization time (the time taken for the eye to fixate on a new location after stimulus), as features for a predictive model. Using data from nine subjects, they reported a best average false acceptance rate of 1.48% achieved with a k-nearest neighbor classifier (k=3).

Subsequently, researchers explored various eye-movement measures including: gaze trajectory [8,11], gaze velocity [46], and pupil size [3] with reasonable success. Galdi et al. developed a gaze analysis (GAS) soft-biometric based on user behavior during observation of particular objects such as facial images [11]. The GAS system constructs a user profile using a fixed area of interest-based feature vector, which is computed using the order-independent cumulative duration of fixations on the respective area of interest. The system was tested on 88 subjects and gave encouraging results on user identification by computing the profile with the lowest Euclidean distance from the test sample.

Yoon et al. explored gaze as a biometric by examining the scanpath of 12 subjects viewing 50 images of patterns with varied spatial characteristics. They modeled gaze velocity using Hidden Markov Models to create unique profiles for each subject. Using a leave-one-out cross-validation scheme, they reported an average performance accuracy in user identification ranging between 53% and 76% [46].

Holland and Oleg evaluated eye movement-based metrics as a feature for biometric identification. They recorded eye-movements while subjects performed a challenging reading task. From the recorded data, they extracted eye-tracking features and scanpath measures including: fixation count, fixation duration, saccade amplitude and velocity. Applying an information fusion method, they combined these features and reported a 27% error rate in a subject identification task [18].

3 Materials and Methods

3.1 Image dataset

For the proposed study, we selected 100 screen-film mammograms from a corpus of mammographic images, digitized using a high resolution LUMISYS scanner (50m per pixel, 12 bit), sourced from the University of South Floridas Digital Database for Screening Mammography (DDSM) [17]. Each case provided by the DDSM database is accompanied by associated patient information, the cranio-caudal (CC) and the mediolateral oblique (MLO) view mammographic images of both the left and the right breasts. Abnormal cases are accompanied by duplicate images containing pixel level ground truth markings of abnormalities, and ground truth subtlety values using the BI-RADS^{TM} lexicon [37] established via biopsy, additional imaging, or two-year follow-up. The selected set included clinically actionable cases covering a broad range of mass margin and shape characteristics. Of the 100 selected cases, 50 cases included biopsy-proven malignant masses, 25 cases included biopsy-proven benign masses, and the remaining 25 cases were normal as determined during a 2-year cancer-free follow-up patient evaluation. A description of the images used in our experiments are provided in greater detail in a previous publication [1].

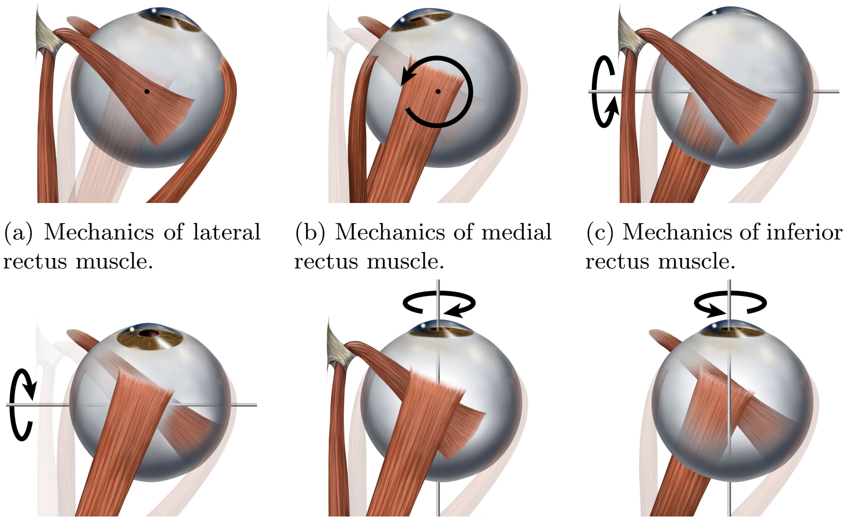
3.2 Experimental Procedure

Ten readers with varied levels of expertise (Radiology residents and board certified radiologists) were recruited from an academic institution to conduct a blind review of the selected mammograms for this study. Each reader was outfitted with an H6 headmounted eye-tracking device developed by Applied Science Laboratories (ASL, Bedford, MA, USA). Readers were then presented with the selected mammographic images on medical grade monitors (dual-head 5MP mammo-grade Totoku LCD monitors calibrated to the DICOM display standard), and asked to report on location and provide a corresponding BI-RADS^{TM} rating of any suspicious mass through a graphical user interface (GUI) custom designed for this experiment. A more detailed overview of the study participants, software and hardware, and the experimental protocol is provided in greater detail in a previous publication [1].

3.3 Eye-Movement Detection

Eye-movements refer to voluntary and involuntary change in the configuration of the eyes, which help the subject to acquire, fixate or track visual stimuli. The movement of the human eye is controlled by pairs of muscles, who's combined and coordinated effect (depicted in Figure 1) is responsible for horizontal (yaw), vertical (pitch), and torsional (roll) eye-movements, respectively; enabling them to control the three-dimensional orientation of the eye.

Three antagonistic pairs of muscles: the lateral and medial rectus muscles, the superior and inferior rectus muscles, and the superior and inferior oblique



(a) Mechanics of lateral rectus muscle.
(b) Mechanics of medial rectus muscle.
(c) Mechanics of inferior rectus muscle.
(d) Mechanics of superior rectus muscle.
(e) Mechanics of superior oblique muscle.
(f) Mechanics of inferior oblique muscle.

Fig. 1: Superior view of muscles responsible for horizontal (yaw), vertical (pitch), and torsional (roll) eye-movements (From Lynch [28]).

muscles, are responsible for the characteristic eye-movements (illustrated in Figure 1) along different axes: horizontal adduction toward the nose or abduction away from it, vertical elevation or depression, and intorsion or extorsion movements that bring the top of the eye toward or away from the nose respectively.

According to Donders law [44], orientation uniquely determines the direction of gaze independent of how the eye was previously orientated. Large sections of the brain control the eye muscles to direct gaze to the desired location in space. Humans primarily engage in seven types of voluntary and involuntary eye-movement: fixation, saccade, glissade, smooth pursuit, microsaccade, tremor, and drift [19]. From eye-tracking data recorded from each reader while reviewing the four mammographic images across two monitors, we extracted fixations and saccades.

A fixation refers to a state in which the eyes remain relatively still (or within a minute spatial radius) over a period of time, such as when the eyes pause on a given word while reading text. The rapid motion of the eye from one fixation to another (such as from one word to another while reading text) is known as a *saccade*. Saccades are considered the fastest movement the body can produce; typically taking 3080 ms to complete. An important peculiarity of saccades is that they rarely take the shortest path between two points, but instead undergo one of several (often suboptimal) paths resulting in *shapes* and *curvatures* 2.

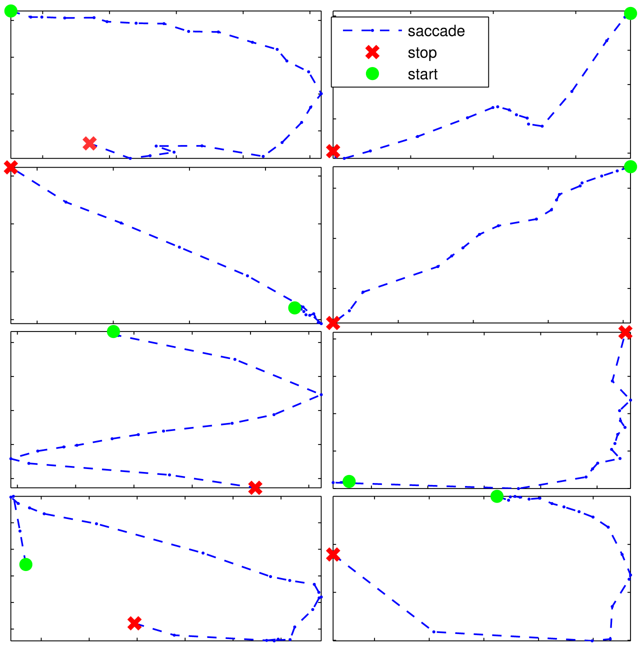


Fig. 2: Sample saccade recorded during a mammographic reading.

Although there is no universally excepted method for computing fixations, there are parameters based on eye physiology, which permit a reasonable criteria for approximating fixations from gaze data. To identify fixations, we computed the average x and y coordinates for gaze points measured over a period of time during which the point-of-gaze continuously remains within an area (approximately 1° visual angle) for a minimum amount of time (approximately 100ms for our algorithm). Since saccades are described in terms of the gaze data between fixations, we computed saccadic events as gaze points connecting the completion of one fixation to the beginning of the next fixation. Saccadic movements between displays (jumping from one screen to the other), thus between mammographic image views, were excluded from our analysis.

3.4 Gesture-based and Geometry-based Features

Once fixation and saccadic events were computed, we applied feature extraction algorithms developed for sketch recognition to characterize the *shape* and *curvature* of individual saccadic movements. Since gaze scanpath is an aggregate shape consisting of individual saccadic movements, aggregating features extracted from saccadic movements will, in principle, provide an accurate characterization of the scanpath.

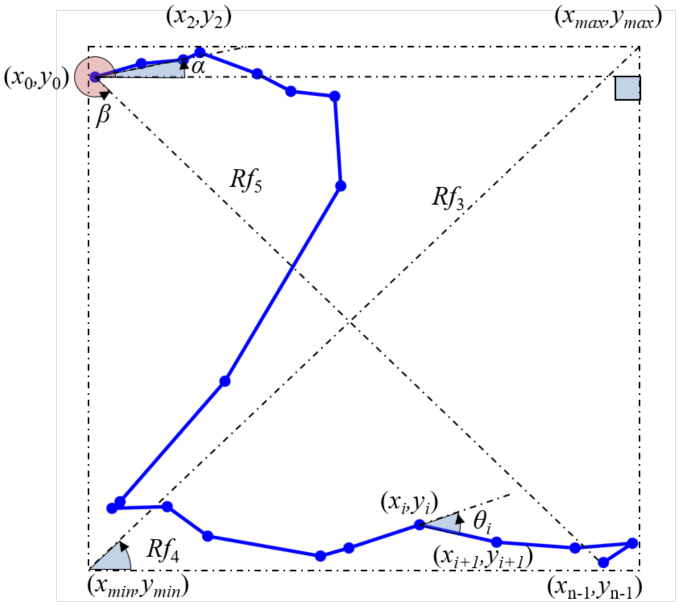


Fig. 3: Rubine's features capture properties associated with the shape a sample saccade from mammographic reading.

Gesture-based features are dependent on how individual strokes are drawn (i.e. the path of each stroke) in contrast to the final geometric shape of the stroke, although the latter can be correlated. For this reason, gesture-based features contain subtle user-dependent variations, which are useful in differentiating between users [10]. Based on work by Rubine [40], Long et al. [27], and Paulson et al. [35], we extracted 29 gesture-based and vision-based features, which were previously demonstrated as being efficiently computable in real-time given a large input size, robust to noise, and capable of encoding semantically meaningful and discriminative information about shapes.

Drawing inspiration from work by Ouyang and Davis [32], we computed an orientation based feature, which captures the direction of the scanpath. The intuition behind this feature is the tendency of the readers' gaze scanpath to follow a specific direction indicating individual behavioral adaptations resulting in a preferred direction for scanning an image. This value is computed as an aggregate of point to point directionality of constituent gaze points in a saccade mapped to one of 12 angles indicating the cardinal direction.

4 Analysis and Results

In this section, we present performance results of sketch-based eye-movement features on two tasks: predicting reader identity, and predicting reader exper-

Table 1: Final feature subsets.

No.	Source	Feature	Description
s8	Sketch	Length of gesture	Length of saccade
s10	Sketch	Angle of gesture	Curvature of saccade
s12	Sketch	Gesture duration	Duration of saccade
s17	Sketch	Linear efficiency	Ratio of saccade length to pixel-wise distance
s18	Sketch	Spatial efficiency	Ratio of saccade length to area covered
s30	Sketch	Gesture orientation	Orientation of saccade
f1	Eye	Pupil size	Average pupil size
f2	Eye	Inter-fixation duration	Time between fixations
f3	Eye	Fixation duration	Duration of fixation
f4	Eye	Scanpath length	Length of scanpath
f5	Eye	Inter-fixation degree	Visual angle between fixations
f6	Eye	No. of fixations	Total No. of fixations
f7	Eye	Fixation rate	Rate of fixations

tise. For comparison purposes, we examined the performance results of traditional eye-tracking features on the same set of tasks. For both sets of features, we first performed feature subset selection to reduce dimensionality of feature representation.

First, since the dependent variable (reader identity identity) is nominal, features were ranked using a combination of model-based, information gain ratio-based, and correlation-based ranking. To compute the model-based ranking, a k-nearest neighbor classifier was trained (one per feature) on a randomly selected training and test subset to predict the identity of each reader.

Information gain (IG) measures the expected reduction in entropy resulting from a partitioning of a dataset based on the values of a given feature. However, IG is not normalized and can therefore be biased in favor of large-valued features. For this reason, we employ the information gain ratio to obtain a gain ratio-based rank for each feature. The information gain ratio (IGR) resolves the limitations of IG by taking the number and size of partitions into account when choosing an attribute, thereby reducing bias towards large-valued attributes.

Next, the ten highest ranked features were selected by combining the gain ratio-based and model-based ranking methods. The final feature set was further reduced by eliminating highly correlated features. Table 1 provides the final subset of features from both the sketch-based features and the traditional eye-tracking features.

We then evaluated the efficacy of both feature subsets by training a Random Forest classifier [4] using a k-fold cross-validation scheme ($k = 10$). For each fold, a 90% of the cases were set aside for training the model, and the remainder 10% was utilized for model evaluation. Note that for each fold, identical cases (identified by case id) were selected from each reader for model training and evaluation. The aggregated (mean) predictive value over all k folds served as the final performance evaluation for the predictive model. All training and testing

Table 2: Performance metrics (*F-score*) for sketch-based and traditional eye-tracking features for biometric identification task.

Reader	**Sketch**	**Eye-tracking**	**ZeroR**
N1	0.94	0.75	0.1
N2	0.88	0.74	0.1
N3	0.9	0.66	0.1
A1	0.87	0.59	0.1
A2	0.87	0.65	0.1
A3	0.92	0.64	0.1
A4	0.95	0.86	0.1
E1	0.87	0.8	0.1
E2	0.88	0.7	0.1
E3	0.84	0.62	0.1
Avg.	0.89	0.7	0.1

evaluations were performed using $WEKA$ software package [12]; an open source machine learning software for building and testing predictive models.

4.1 Predicting Reader Identity

To test the effectiveness of sketch-based features on a biometric identification, we developed a between-subject predictive model using a Random Forest classifier evaluated using a k-fold cross-validation partitioning scheme ($k = 10$) as previously described. Multiple (k) rounds of cross-validation were performed using different partitions, and the validation results were averaged over all rounds in to reduce variability. As a baseline, we include the results of a majority classifier (ZeroR). A ZeroR classifier is a simple majority rule classifier, which classifies all input test samples as the majority or modal class independent of feature values of the input sample. In Table 2, we report *F-score* (the harmonic mean of precision and recall) performance metrics for the biometric identification task

Table 3: Confusion matrix for sketch-based features for biometric identification task.

		PREDICTED									
		NR1	NR2	NR3	AR1	AR2	AR3	AR4	E1	E2	E3
ACTUAL	NR1	93	1	0	2	0	1	0	0	0	3
	NR2	0	90	1	1	1	0	3	4	0	0
	NR3	2	0	91	3	0	0	3	0	0	1
	AR1	1	2	0	94	0	0	0	0	1	2
	AR2	2	0	1	2	83	3	0	2	2	5
	AR3	4	3	0	1	3	89	0	0	0	0
	AR4	0	1	0	2	0	0	96	1	0	0
	E1	1	2	2	0	0	0	0	85	7	3
	E2	0	1	1	3	2	0	0	8	85	0
	E3	3	1	3	4	1	0	0	2	3	83

Table 4: Performance metrics (*F-score*) for sketch-based and eye-tracking features for reader expertise prediction task.

Class	**Sketch**	**Eye-tracking**	**ZeroR**
NR	0.9	0.77	0.4
AR	0.93	0.8	0.4
E	0.91	0.83	0.4
Avg.	0.91	0.8	0.4

using sketch-based features. For comparison purposes, Table 2 also includes performance metrics using eye-tracking features for the same task. The confusion matrix provided in Table 3 illustrates the instances of error when predicting the actual class label for the sketch-based models.

4.2 Predicting Reader Expertise

We grouped each of the 10 participating readers into one of three experience levels: new trainee resident (NR), advanced trainee resident (AR), and expert radiologist (E). Next, utilizing a similar cross-validation partitioning scheme, we evaluated the efficacy of sketch-based features in predicting the experience level (expertise) of each reader. In Table 4,we report *F-score* performance metrics for the reader expertise prediction task using sketch-based features and include the performance of eye-tracking features for the same task for comparison. The confusion matrix provided in Table 5 illustrates the instances of error when predicting the actual class label for the sketch-based models.

5 Discussion

The final set of features (see Table 1) include four measures related to motion: the orientation, duration, length, and rotational change of the shape formed by the saccade, and two measures of visual appearance: ratio of saccade length to overall size (*s16*), and the ratio of saccade length to the actual inter-fixation distance (*s17*). The highest ranked feature, saccade orientation, explains the tendency of the image readers' saccadic scanpath to follow a specific direction. We speculate that this feature captures coordinated muscle movements resulting from adaptations of repetitive behavior over time, which are specific to the

Table 5: Confusion matrix of predictive model for reader expertise using sketch-based features from eye-movement.

		PREDICTED		
ACTUAL		NR	AR	E
	NR	270	17	13
	AR	15	370	15
	E	15	13	272

individual. This observation is not unlike the uniqueness of the biomechanics of walking (gait). However, more detailed studies and experimental data is required to validate this speculative statement.

Previous studies in mammography have identified some measures of direction, duration, and lengths of saccadic movements as containing discriminative information about the experience in radiology [25,30,24]. Intuitively, both density metrics (*s16* and *s17*) capture the spatial efficiency of the saccadic movements. While *s16* measures the linear efficiency of the scanpath, *s17* measures the two-dimensional spatial efficiency of the scanpath. Both features give a piecewise decomposition of the geometric properties of the scanpath formed by the image reader during the screening process. Previous studies have suggested that measures of overall scanpath formed during the viewing of a mammographic case are related to the individual and experience [25,31]. The scanpath has also been studied as a biometric for individual identification under varied image viewing conditions unrelated to mammography [46,18]. To the best of our knowledge, these features were never applied in predictive models as biometric identifiers or for predicting experience level. Additionally, the characterization using gesture recognition methods have never been explored until now.

6 Conclusions

In this study, we proposed and evaluated two methods for extracting features, which contain discriminative information about the identity and the level of expertise of a radiologist in screening mammography. These features characterize changes in positional and non-positional measures of eye-movement. First, we applied sketch recognition algorithms to extract gesture and geometry-based features from eye-tracking data. These features give a fine-grained characterization of the scanpath by aggregating the spatial (shape), directional, and kinetic properties of individual saccadic movements. We compared the effectiveness of these sketch-based features with more traditional metrics from eye-tracking.

Using a corpus of eye-movement and pupillary data from 100 mammographic cases reviewed by ten readers of varied experience level, recorded under clinically equivalent experimental conditions, the findings presented in this study establishes the following generalizable trends:

1. During the mammographic screening task, positional and non-positional measures of changes in the eye can provide sufficient discriminative information about the identity of an image reader.
2. Positional and non-positional measures of changes in the eye provide sufficient discriminative characterization of the readers' level of expertise for a given task (mammographic screening).
3. Both positional and non-positional measures perform significantly better than random chance at predicting the readers' identity and level of expertise.
4. Sketch-based features of eye-movement result in more accurate predictive models when compared with more traditional eye-tracking features.

7 Acknowledgement

The authors thank Kathleen B. Hudson, MD, and Garnetta Morin-Ducote, MD, for their invaluable contributions to this work.

References

1. Alamudun, F., Yoon, H.J., Hudson, K.B., Morin-Ducote, G., Hammond, T., Tourassi, G.D.: Fractal analysis of visual search activity for mass detection during mammographic screening. Medical Physics pp. n/a–n/a (2017), http://dx.doi.org/10.1002/mp.12100
2. Bartley, J., Forsyth, J., Pendse, P., Xin, D., Brown, G., Hagseth, P., Agrawal, A., Goldberg, D.W., Hammond, T.: World of workout: a contextual mobile rpg to encourage long term fitness. In: Proceedings of the 2nd ACM SIGSPATIAL International Workshop on the Use of GIS in Public Health. pp. 60–67. ACM (2013)
3. Bednarik, R., Kinnunen, T., Mihaila, A., Fränti, P.: Eye-movements as a biometric. In: Image Analysis, pp. 780–789. Springer (2005)
4. Breiman, L.: Random forests. Machine learning 45(1), 5–32 (2001)
5. Chi, M.T., Glaser, R., Farr, M.J.: The nature of expertise. Psychology Press (2014)
6. Çığ, Ç., Sezgin, T.M.: Gaze-based prediction of pen-based virtual interaction tasks. International Journal of Human-Computer Studies 73, 91 – 106 (2015), http://www.sciencedirect.com/science/article/pii/S1071581914001244
7. Corey, P., Hammond, T.: Gladder: combining gesture and geometric sketch recognition. In: Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 3. pp. 1788–1789. AAAI Press (2008)
8. Deravi, F., Guness, S.P.: Gaze trajectory as a biometric modality. In: BIOSIGNALS. pp. 335–341 (2011)
9. Dixon, D., Prasad, M., Hammond, T.: icandraw: using sketch recognition and corrective feedback to assist a user in drawing human faces. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 897–906. ACM (2010)
10. Eoff, B.D., Hammond, T.: Who dotted that'i'?: context free user differentiation through pressure and tilt pen data. In: Proceedings of Graphics Interface 2009. pp. 149–156. Canadian Information Processing Society (2009)
11. Galdi, C., Nappi, M., Riccio, D., Cantoni, V., Porta, M.: A new gaze analysis based soft-biometric. In: MCPR. pp. 136–144. Springer (2013)
12. Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The weka data mining software: an update. ACM SIGKDD Explorations Newsletter 11(1), 10–18 (2009)
13. Hammond, T., Gajos, K., Davis, R., Shrobe, H.E.: An agent-based system for capturing and indexing software design meetings. In: In Proceedings of the International Workshop on Agents in Design (WAID02. Citeseer (2002)
14. Hammond, T., OSullivan, B.: Recognizing freeform hand-sketched constraint network diagrams by combining geometry and context. In: Proceedings of the Eurographics Ireland (2007)
15. Hammond, T., Prasad, M., Dixon, D.: Art 101: learning to draw through sketch recognition. In: Smart Graphics. pp. 277–280. Springer (2010)

16. Hammond, T.A.: Ladder: A perceptually-based language to simplify sketch recognition user interface development. Ph.D. thesis, Massachusetts Institute of Technology (2007)
17. Heath, M., Bowyer, K., Kopans, D., Kegelmeyer Jr, P., Moore, R., Chang, K., Munishkumaran, S.: Current status of the digital database for screening mammography. In: Digital mammography, pp. 457–460. Springer (1998)
18. Holland, C., Komogortsev, O.: Biometric identification via eye movement scanpaths in reading. In: Biometrics (IJCB), 2011 International Joint Conference on. pp. 1–8 (Oct 2011)
19. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., Van de Weijer, J.: Eye tracking: A comprehensive guide to methods and measures. Oxford University Press (2011)
20. Kara, L.B., Stahovich, T.F.: An image-based, trainable symbol recognizer for hand-drawn sketches. Computers & Graphics 29(4), 501 – 517 (2005), `http://www.sciencedirect.com/science/article/pii/S0097849305000853`
21. Kasprowski, P., Ober, J.: Biometric Authentication: ECCV 2004 International Workshop, BioAW 2004, Prague, Czech Republic, May 15th, 2004. Proceedings, chap. Eye Movements in Biometrics, pp. 248–258. Springer Berlin Heidelberg, Berlin, Heidelberg (2004), `http://dx.doi.org/10.1007/978-3-540-25976-3_23`
22. Krupinski, E.A.: Influence of experience on scanning strategies in mammography. In: Proc SPIE Medical Imaging. pp. 2712:95–101. International Society for Optics and Photonics (1996)
23. Krupinski, E.A., Graham, A.R., Weinstein, R.S.: Characterizing the development of visual search expertise in pathology residents viewing whole slide images. Human pathology 44(3), 357–364 (2013)
24. Krupinski, E.A., Tillack, A.A., Richter, L., Henderson, J.T., Bhattacharyya, A.K., Scott, K.M., Graham, A.R., Descour, M.R., Davis, J.R., Weinstein, R.S.: Eye-movement study and human performance using telepathology virtual slides. implications for medical education and differences with experience. Human Pathology 37(12), 1543 – 1556 (2006), `http://www.sciencedirect.com/science/article/pii/S0046817706005302`
25. Kundel, H.L., Paul S. La Follette, J.: Visual search patterns and experience with radiological images. Radiology 103(3), 523–528 (1972), `http://dx.doi.org/10.1148/103.3.523`, pMID: 5022947
26. Li, W., Hammond, T.: Recognizing text through sound alone. In: Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence. pp. 1481–1486. AAAI'11, AAAI Press (2011), `http://dl.acm.org/citation.cfm?id=2900423.2900657`
27. Long, Jr., A.C., Landay, J.A., Rowe, L.A., Michiels, J.: Visual similarity of pen gestures. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 360–367. CHI '00, ACM, New York, NY, USA (2000), `http://doi.acm.org/10.1145/332040.332458`
28. Lynch, P.J.: Eye movements abductors. `https://en.wikipedia.org/wiki/File:Eye_movements_abductors.jpg` (Dec 2006)
29. Miller, J., Hammond, T.: Wiiolin: a virtual instrument using the wii remote. In: Proceedings of the 2010 Conference on New Interfaces for Musical Expression (NIME). pp. 497–500 (2010)
30. Nodine, C.F., Kundel, H.L., Mello-Thoms, C., Weinstein, S.P., Orel, S.G., Sullivan, D.C., Conant, E.F.: How experience and training influence mammography expertise. Academic Radiology 6(10), 575 – 585 (1999), `http://www.sciencedirect.com/science/article/pii/S1076633299802529`

31. Noton, D., Stark, L.: Scanpaths in eye movements during pattern perception. Science 171(3968), 308–311 (1971), http://science.sciencemag.org/content/171/3968/308
32. Ouyang, T.Y., Davis, R.: A visual approach to sketched symbol recognition. In: Proceedings of the 21st International Joint Conference on Artificial Intelligence. pp. 1463–1468. IJCAI'09, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA (2009), http://dl.acm.org/citation.cfm?id=1661445.1661680
33. Paulson, B., Cummings, D., Hammond, T.: Object interaction detection using hand posture cues in an office setting. International Journal of Human-Computer Studies 69(1), 19–29 (2011)
34. Paulson, B., Hammond, T.: Paleosketch: Accurate primitive sketch recognition and beautification. In: Proceedings of the 13th International Conference on Intelligent User Interfaces. pp. 1–10. IUI '08, ACM, New York, NY, USA (2008), http://doi.acm.org/10.1145/1378773.1378775
35. Paulson, B., Rajan, P., Davalos, P., Gutierrez-Osuna, R., Hammond, T.: What!?! no rubine features?: Using geometric-based features to produce normalized confidence values for sketch recognition. In: HCC Workshop: Sketch Tools for Diagramming (VL/HCC). pp. 57–63 (2008)
36. Pramanik, S., Bhattacharjee, D.: Geometric feature based face-sketch recognition. In: Pattern Recognition, Informatics and Medical Engineering (PRIME), 2012 International Conference on. pp. 409–415 (March 2012)
37. of Radiology. BI-RADS Committee, A.C., of Radiology, A.C.: Breast imaging reporting and data system. American College of Radiology (1998)
38. Rajan, P., Hammond, T.: From paper to machine: extracting strokes from images for use in sketch recognition. In: Proceedings of the 5th Eurographics Conference on Sketch-Based Interfaces and Modeling. pp. 41–48. Eurographics Association (2008)
39. Rayner, K.: Eye movements in reading and information processing: 20 years of research. Psychological Bulletin 124(3), 372 (1998)
40. Rubine, D.: Specifying gestures by example. SIGGRAPH Comput. Graph. 25(4), 329–337 (Jul 1991), http://doi.acm.org/10.1145/127719.122753
41. Schnitzer, B.S., Kowler, E.: Eye movements during multiple readings of the same text. Vision Research 46(10), 1611 – 1632 (2006), http://www.sciencedirect.com/science/article/pii/S0042698905004864
42. Smith, R.A., Cokkinides, V., von Eschenbach, A.C., Levin, B., Cohen, C., Runowicz, C.D., Sener, S., Saslow, D., Eyre, H.J.: American cancer society guidelines for the early detection of cancer. CA: A Cancer Journal for Clinicians 52(1), 8–22 (2002)
43. Taele, P., Peschel, J., Hammond, T.: A sketch interactive approach to computer-assisted biology instruction. In: Proc. IUI 2009 Workshop on Sketch Recognition (2009)
44. Tweed, D., Vilis, T.: Geometric relations of eye position and velocity vectors during saccades. Vision Research 30(1), 111 – 127 (1990), http://www.sciencedirect.com/science/article/pii/0042698990901314
45. Valentine, S., Vides, F., Lucchese, G., Turner, D., Kim, H.h., Li, W., Linsey, J., Hammond, T.: Mechanix: A sketch-based tutoring system for statics courses. In: Proceedings of the 24th Innovative Applications of Artificial Intelligence Conference on Artificial Intelligence (IAAI). pp. 2253–2260. AAAI (2012)
46. Yoon, H.J., Carmichael, T.R., Tourassi, G.: Gaze as a biometric. In: Proc. SPIE. pp. 903707–903707 (2014)