

COSC460

Improving Face Recognition with Genealogical and Contextual Data

November 2, 2012

Ellie Rasmus

`emr51@uclive.ac.nz`

Supervisor: Dr. Richard Green

`richard.green@canterbury.ac.nz`

Abstract

Face recognition has long been an area of great interest within computer science, and as face recognition implementations become more sophisticated, the scope of real-world applications has widened. The field of genealogy has embraced the move towards digitisation, with increasingly large quantities of historical photographs being digitised in an effort to both preserve and share them with a wider audience. Genealogy software is prevalent, but while many programs support photograph management, only one uses face recognition to assist in the identification and tagging of individuals. Genealogy is in the unique position of possessing a rich source of context in the form of a family tree, that a face recognition engine can draw information from. We aim to improve the accuracy of face recognition results within a family photograph album through the use of a filter that uses available information from a given family tree. We also use measures of co-occurrence, recurrence and relative physical distance of individuals within photos to accurately predict their identities. This proposed use of genealogical and contextual data has shown a 26% improvement in accuracy over the most advanced face recognition technology currently available when identifying 348 faces against a database of 523 faces. These faces are extracted from a challenging dataset of 173 family photographs, dating back as far as 1908.

Acknowledgments

I'd like to thank Glen Cameron from NEC NZ for his assistance in setting up and using Neo-Face, Janice Newcombe for her help gathering the photographs and genealogical data needed for this research, Christoph Bartneck from the HIT Lab NZ, who initiated this research, and Richard, for being a great supervisor.

Contents

1	Introduction	4
2	Background and Related Work	5
2.1	Face Detection and Recognition	5
2.1.1	Existing Use within Genealogy	5
2.1.2	Challenges Applying Face Recognition to Historical Photographs	7
	Uneven lighting	7
	Variation of pose	8
2.1.3	Approaches to Face Recognition	8
	Holistic Methods	8
	Feature-based Methods	8
	Hybrid Methods	9
	POEM Descriptors	9
2.1.4	NeoFace Recognition Engine	9
	Face Detection	9
	Face Recognition	9
2.2	Use of Contextual Information and Metadata	10
2.2.1	Digital Metadata	10
2.2.2	Estimating Relationships	11
	Physical Measures of Context	11
	Recurrence and Co-occurrence	11
	Filtering out impossibilities	12
2.3	Genealogical Data	12
2.3.1	Degree of familial relationship	12
2.3.2	The GEDCOM standard	12
3	Design and implementation	14
3.1	Overview	14
3.2	GEDCOM Files	14
3.3	Face Recognition	14
3.3.1	Face Grouping with NeoFace	14
	Matching Threshold	14
3.3.2	Age and Gender Estimation	15
3.4	Context Filter	16
3.4.1	Time Frames	16
	Photograph Time Frames	17
	Face Group Time Frames	17
	Candidate lists	17
3.4.2	Degree of Familial Relationship	18
	Calculating the Coefficient of Relationship	18
	Using the Coefficient of Relationship	19
3.4.3	Co-occurrence	20
	Number of people in photograph	20
	Physical Distance	20
3.4.4	Recurrence	20
3.5	System Structure	21
3.6	Composing the final score	21

4 Results	24
4.1 Experimental Design	24
4.2 Overall Accuracy	24
4.3 Accuracy vs. Identity Recurrence	25
4.4 Accuracy vs. Candidate Co-occurrence	25
4.5 Accuracy vs. Relatedness To Photo	27
4.6 Accuracy vs. Candidate List Size	27
4.7 Accuracy of Young Face Identification	28
4.8 Discussion	29
5 Conclusion and Future Work	31

1

Introduction

The field of genealogy has embraced the move towards digitisation, with increasingly large quantities of historical photographs being digitised in an effort to both preserve and share with a wider audience. Genealogy software is prevalent, but while many programs support photograph organisation and management, few allow for the tagging of individuals in photos, and only one uses face recognition to assist in the process of face identification and tagging. Genealogy is in the unique position of possessing a rich source of context, in the form of a family tree, that a face recognition engine can draw information from.

Unlike standard collections of modern photos, family photo albums usually span over a much longer period of time, and often contain very old, poor quality photos that face recognition technology struggles with.

Also, face recognition often suffers from poor performance when identifying the faces of children and young people. Photos of young individuals often appear very different to when they are adult, and the faces of children are often incorrectly matched to faces of other children.

The aim of this research is to improve the speed and efficiency with which a user can identify and tag individuals in a large collection of family photographs by using contextual information from a provided family tree to estimate relationships. Alongside this, we also derive contextual measures of co-occurrence, recurrence and relative physical distance of individuals appearing in photos within the album. We use these genealogical and contextual measures alongside face recognition to accurately predict the identity of an unknown face.

2

Background and Related Work

2.1 Face Detection and Recognition

2.1.1 Existing Use within Genealogy

While some photo management applications such as Google’s Picasa make use of face recognition technology to tag faces within photo albums, these are intended for modern photo collections, and have no facilities for integrating genealogical information [6].

Genealogical Software Many genealogical software programs provide photograph and media management facilities. However, this rarely includes face tagging, and only one genealogical software program makes use of face recognition.

Gramps is an open-source genealogical software project that provides some facilities for photo management [3]. However, these are not extensive, and the photo album is not integrated at all with the family tree; the photos are isolated in their own folder, with no connection made between the individuals in the photos and the individuals in the family tree (see Fig. 2.1).

Family Tree Maker, released by the very popular Ancestry.com, is probably the most widely used desktop genealogy application [2]. The integration of media improves on that in Gramps, with connections being made between individuals in photos and the family tree itself. However, no actual face tagging occurs; we know who appears in a photo, but not exactly where they are (see Fig. 2.2). Any additional information as to the location of individuals is maintained in a ‘note’ alongside the photo.

The only genealogical web service available that uses any face recognition technology is MyHeritage [5], where detected faces can be matched against a celebrity face database to find famous lookalikes. However this only allows for matching against a predefined set of celebrity faces, not a user-selected set of family photos. While faces are automatically detected, tagging of faces in family photos is still a manual task. During the course of this research, MyHeritage has added face recognition technology to their desktop genealogy application Family Tree Builder.¹ [4] This is currently the only genealogy software that uses face recognition to identify and tag faces. It does not, however, make use of any additional context in the face identification process.

Measure of Familial Similarity The KinFace database is a set of photographs compiled by the University at Buffalo, New York, to enable the verification of familial likeness measures provided by recognition algorithms [43, 41, 30]. In [40, 30], a method of estimating whether

¹This is not to be confused with Ancestry.com’s Family Tree Maker - many websites appear to confuse the two

2.1. FACE DETECTION AND RECOGNITION

6

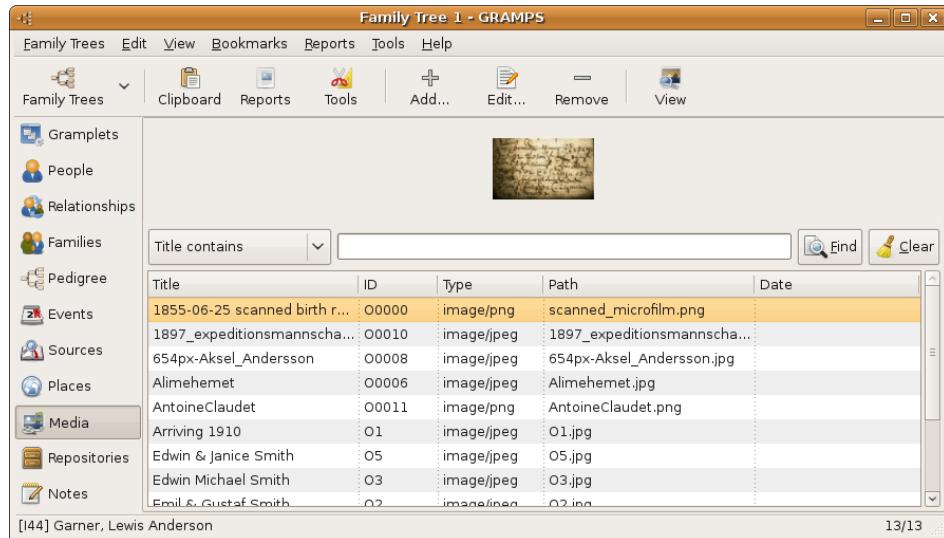


Figure 2.1: Photo management interface in Gramps.

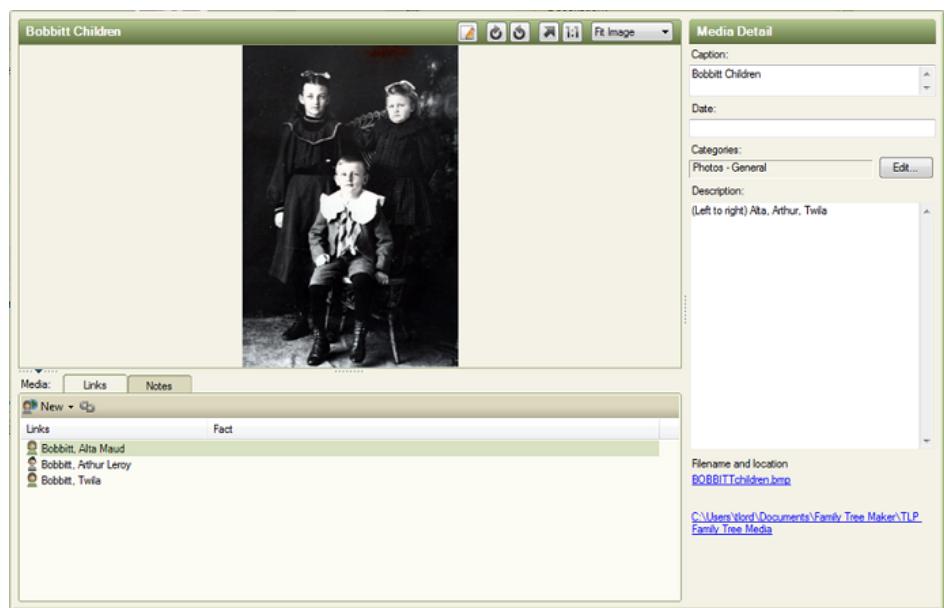


Figure 2.2: Photo management interface in Ancestry.com's Family Tree Maker.

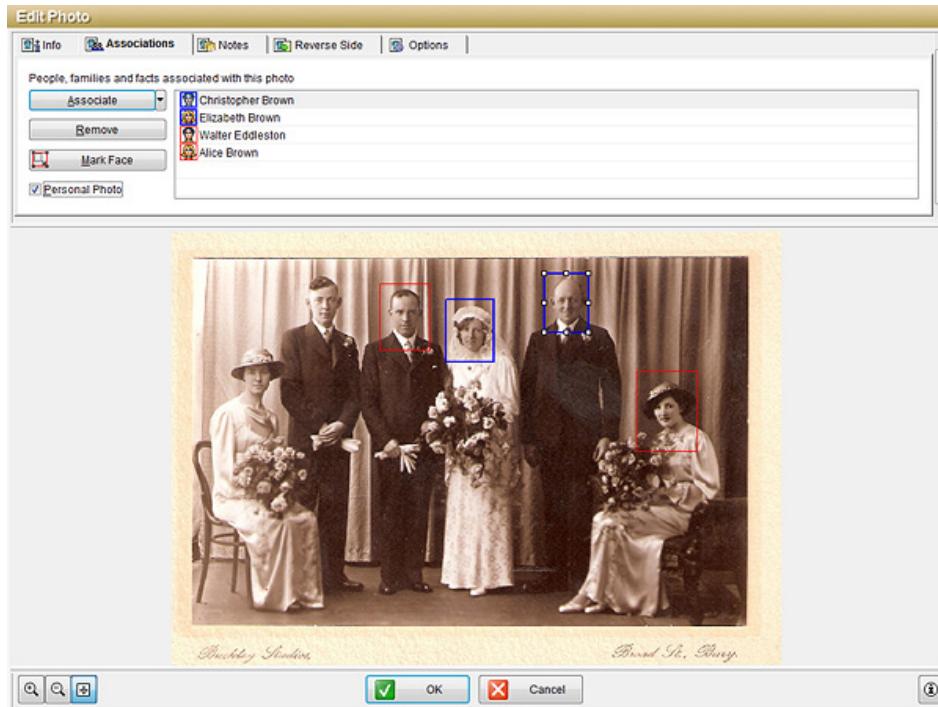


Figure 2.3: Face tagging interface in MyHeritage’s Family Tree Builder.

two individuals are related is introduced, through the extraction of genetic-invariant facial features. Facial structure information is also used, derived from 6 key facial points. The ultimate goal of this research is to prove some genetic relationship between two individuals through analysis of photographs alone, rather than pairing a unknown face’s identity to a known one.

2.1.2 Challenges Applying Face Recognition to Historical Photographs

In order to gain accurate results from existing face recognition implementations, the faces being studied often need to be straight on to the camera, evenly lit and fully visible. Unfortunately, this is rarely the case, and historical photos introduce the additional problem of poor image quality caused by inferior technology and degradation over time.

Uneven lighting

Yan et al. categorise previous attempts to combat illumination variation in photos into three groups: invariant feature extraction, face modelling, and pre-processing and normalisation [54]. Invariant feature extraction is typified by Shashua et al.’s work on quotient images [44]. This approach involves a series of subtractions, creating a face image consisting of areas not affected by light variation, which is then mapped onto a 3D face model. However, this method requires that face images be well-aligned, and assumes that the 3D face model fits all potential faces universally. An example of the face modelling approach is the illumination cone model, which uses sample images of a subject in a single pose under varying lighting to construct a convex cone [20]. This cone can be manipulated to represent all lighting conditions and poses. However, this approach is computationally expensive, and requires seven training

images of each subject in a single pose. An example of the pre-processing and normalisation approach is image manipulation by global histogram equalisation [39, 7, 16]. While this is a computationally inexpensive operation, it does not handle complex varying illumination over an image. Another popular approach to normalisation is by discrete wavelet transform [12, 21]. This normalises the reflection and illumination components of an image through high and low pass filters, resulting in an image affected minimally by any shadows or illumination artefacts present in the original. This is generally successful and computationally inexpensive. An integrated normalisation approach is proposed by Kao et al. that involves local contrast enhancement, where a given pixel is represented by the ratio of intensity of the pixel to the average of its surrounding pixels [26]. A discrete cosine transform is also applied to support feature selection. This approach has had promising results.

Variation of pose

The illumination cone model mentioned above deals not only with lighting variations but also variation of pose, through manipulation of the cone to which the subject's face has been applied [20]. This approach still has the downsides mentioned above. Ana Paula et al. attempted to solve the problem of varying placements of the face within a detected area by creating 12 different versions of the same face image, with the face in a different placement in each one [8]. Each version was then tested against a pre-existing face database, and the image closest to the average face was selected. This, however, is computationally expensive, and does not deal with the issue of different poses, just faces being positioned in different places around the detected area. A 3D face modelling approach is proposed by Xiaozheng et al. [53]; however this requires two face images of set poses, frontal and profile, which are not easily obtained within a set of historical images. Beymer introduces a method of pose estimation where detected facial features are applied to various predefined templates, but again, many training images of each subject are required for success [10].

2.1.3 Approaches to Face Recognition

Zhao et al. categorise popular face recognition attempts into three categories: holistic matching methods, feature-based matching methods and hybrid matching methods [55].

Holistic Methods

These methods use the whole face as input to the recognition system. The most widely accepted technique to develop from the holistic approach is that of Eigenfaces, based on principal component analysis. This approach uses a large set of faces to derive an average face to compare a given face against [48]. This approach, however, relies on uniform pose and illumination, and can result in poor results when input images deviate from this [29].

Feature-based Methods

These methods extract facial features such as eyes, nose and mouth, to use for recognition. A well-known feature-based approach uses hidden Markov models. This involves the location and extraction of localised feature blocks from face images in the form of 2D discrete cosine transform vectors [34]. However, this approach has shown weak results when applied to photos of varying pose and illumination [29].

Hybrid Methods

Hybrid methods attempt to incorporate both holistic and feature-based approaches. The best example of this is the Eigenspace or Modular Eigenface approach, which introduces a layered representation of a face consisting of an Eigenface and a series of Eigenfeatures [37]. This modularisation of the face-space allows for more accurate description of faces and copes better with faces in varying pose.

POEM Descriptors

A new approach to face recognition is based on the use of patterns of oriented edge magnitudes (POEM) [49]. This takes into consideration the distribution of edges in different areas of an image, and applies a local binary pattern operator for classification of features. This provides fast, accurate results; however, as it is a relatively new approach, there has not been a great deal of evaluation to back up these results.

2.1.4 NeoFace Recognition Engine

NeoFace is a face recognition system developed by NEC [42, 24]. In 2010, NeoFace was entered into the Multiple Biometric Evaluation Challenge, an evaluation of face recognition's current state of the art performed by the National Institute of Standards and Technology (NIST) [38]. It achieved an identification rate of 95% on a database of 1.8 million individuals, the highest rate of all participating technologies.

Face Detection

Detection of potential faces is performed using generalized learning vector quantisation; the model is trained with a large number of face and non-face images, and the images in the training set that match the query image the closest are merged together to provide face and facial feature locations [42, 24]. The level of confidence that a detected face is in fact a face is calculated by $p = (d_{nf} - d_f)/(d_{nf} + d_f)$ where d_f is the Euclidean distance between the query image and the nearest feature space template that is in the face category, and d_{nf} is the distance from the closest non-face template. If the query image is closer to a non-face template, the value will be less than 0. If it is more than 0, the query image is considered to be a face [42].

Face Recognition

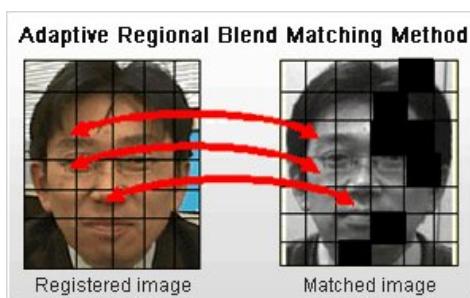


Figure 2.5: Source image divided into regions.

NeoFace attempts to combat both global image variations (such as changes in illumination and pose) and local image variations (such as the appearance of sunglasses, or differing facial expressions). Global image variations are accounted for by applying the perturbation space method, which involves generating many images from one enrolled image [42, 24]. These are generated to simulate different face poses and illumination conditions by using predefined shape and illumination models (see Fig. 2.4).

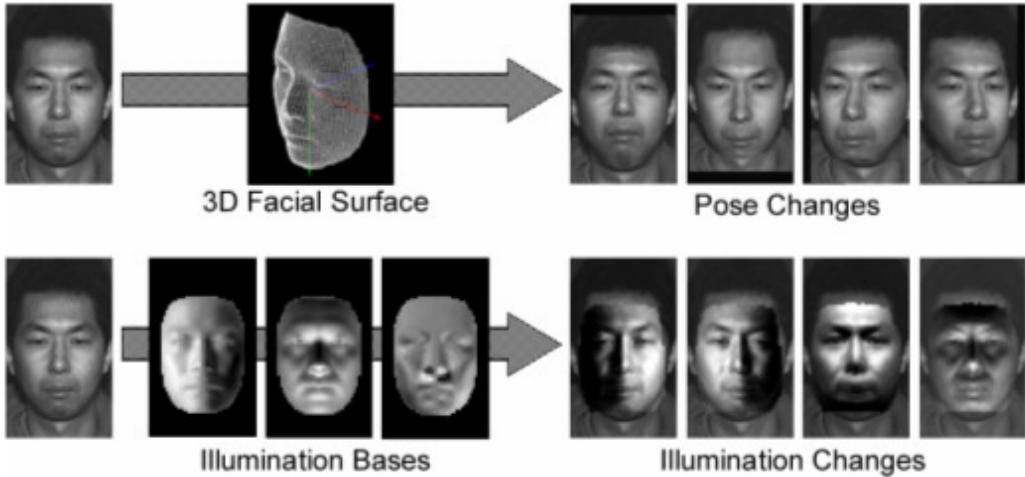


Figure 2.4: Multiple face images to be enrolled after perturbation space method is applied to source image.

Local image variations are accounted for by using adaptive regional blend matching [42]. This involves splitting up both query and target images into a number of regions, and scoring each region in the query image for similarity to the corresponding region in the target image (see Fig. 2.5). Only highly scoring regions are taken into account for final similarity calculation, meaning that any areas of localised difference, such as the appearance of sunglasses or a different facial expression, will not weigh heavily on the final score if there are other regions of high similarity.

NeoFace's approach to face recognition could be considered a hybrid one, with the perturbation space method addressing variation across the whole face, and adaptive regional blend matching dealing with variation of features.

2.2 Use of Contextual Information and Metadata

2.2.1 Digital Metadata

Despite the significant challenges posed by historical photos with respect to image quality and available metadata, genealogy is in the unique position of having additional information to supplement a collection of photographs. This can be used to filter out inappropriate results. While the use of context to improve the efficiency of photo classification has been widely investigated, much of this relies heavily on the availability of extensive metadata supplied by a digital camera at the time of the photograph being taken, such as timestamps and GPS coordinates [33, 31, 25, 28, 13, 18, 14]. Realistically, only a small proportion of a family photograph album will have such information, as many images will have been digitised with a scanner or similar copying device. Much of this research also relies on the photographs in the album being in colour, providing additional visual context within an event from clothing, background, and skin colour [9, 46, 25, 23].

Work by Naaman et al. outlines the contextual information that is most important in assisting user photograph navigation and location [32]. Of the 24 metadata categories tested, the 7 of highest importance were:

- Whether scene is indoors or outdoors
- The people within a photograph
- The location
- The event depicted
- The number of people
- The season
- The year

This study relied on geo-referenced photographs with comprehensive digital metadata. Not all photographs have this kind of information available alongside them; we can still attempt to make use of some of these categories, such as the identities and number of people in a photograph, and an estimated year or time frame of photograph origin.

2.2.2 Estimating Relationships

Physical Measures of Context

Wang et al.'s study into recognising social relationships through photograph analysis uses a number of physical measurements to estimate relationships between unidentified individuals in a photograph [50]. Height difference is used to estimate relationships such as parent-child and husband-wife, where the former will often appear below the latter in a photo due to physical height. The ratio of two face sizes is used to estimate age within potential relationships; it is more likely that two individuals are parent and young child if one face is much smaller than the other. The physical closeness of two people is used as a factor impacting relationship estimation, as is an estimation of the age difference between two individuals. Gender estimation is also used to determine each person's role in a relationship. Relative face size is found to be the most useful measure for relationship estimation. This research also finds that a relationship model can be successfully trained on a number of photographs, then used to predict relationships in a different set of images.

Ng et al. also explore the use of physical distance as a semantic tool [35]. The normalised estimated distance $D(p_i, p_j)$ between two individuals p_i, p_j is calculated by $D(p_i, p_j)/D * (p_i, p_j) = w_s/w(p)$, where $D * (p_i, p_j)$ is the euclidean distance between faces in the photograph, w_s is a constant average head size (of 140mm), and $w(p)$ is the average head size in the photograph.

Ding et al. find that visual concepts derived from video footage can be analysed to benefit the estimation of social relationships [17]. This study is more concerned with determining the valency of a relationship, but does show that the consideration of visual patterns can be used as a form of context when estimating relationships.

Recurrence and Co-occurrence

The concept of using a measure of recurrence of individuals within photos to improve identification has been explored in [25, 33, 46, 52]; however much of their functionality relies on recurrence within a particular event, which requires both timestamps on photos and for many photos to be from one event. This is unlikely in historical family photos, with fewer photos being taken in the past. The use of identity co-occurrence within photo collections has been explored in many studies [45, 15, 11, 52, 27, 35]. Shimizu et al.'s basic use of co-occurrence provides promising results, but requires manual confirmation of photograph relevance [45]. The use of context acquired solely from co-occurrence in photos within a social network, without any visual feature analysis, has provided good accuracy [15, 18, 47]; however these solutions are designed for use within a large scale online social network rather than in a personal collection. It also makes use of a great deal of contextual information that is not

available in a genealogical application, such as who has made comments on a photo. Wu et al. uses multiple social context measures such as co-occurrence, physical distance between faces, and the number of individuals in a photograph, to improve face clustering accuracy [52]. This approach relies on three factors:

- If two faces are closer, their relationship is likely to be closer
- If more people appear in a photograph, the physical distance between faces is likely to be of less importance
- If two individuals appear in the same photographs more often, the distance between their faces is likely to be of higher importance.

This approach was tested on a group of 259 photos from 4 separate events, and provided encouraging results.

Filtering out impossibilities

A simple context filter ensuring that one person is not identified in the same photo twice has been used in face recognition and tagging software Faint [1], but additional genealogical information means that we can rule out many candidates quickly by considering the time the photo was taken as well as the birth and death dates of individuals.

2.3 Genealogical Data

2.3.1 Degree of familial relationship

The coefficient of relationship expresses the consanguinity or theoretical percentage of genes two individuals share, determined by their common ancestors [51]. This can be calculated for two individuals a and b by $R_{ab} = \sum(1/2)^n$, where n is the number of paths separating a and b .

2.3.2 The GEDCOM standard

An acronym for Genealogical Data Communication, GEDCOM files are the current standard for recording and sharing genealogical data [36]. GEDCOM 5.5 is still the most widely used and accepted version, despite being released in December of 1995. GEDCOM files are structured similarly to an XML file, with the number at the start of each line indicating the depth or level of the following data.

Figure 2.6 shows an excerpt from a GEDCOM file. The file structure is based on family (FAM) and individual (INDI) entities. A family entity models a nuclear family, which consists of two parents and their children.

This 17-year-old standard has very limited support for the incorporation of photographs and multimedia objects. Many existing genealogical programs create their own file format, often building on top of GEDCOM's functionality. The problem this poses is the loss of an actual standard; often, only the software that wrote the file will be able to read it again, hampering the open distribution of information, especially if the program necessary to open the file is not freely available.

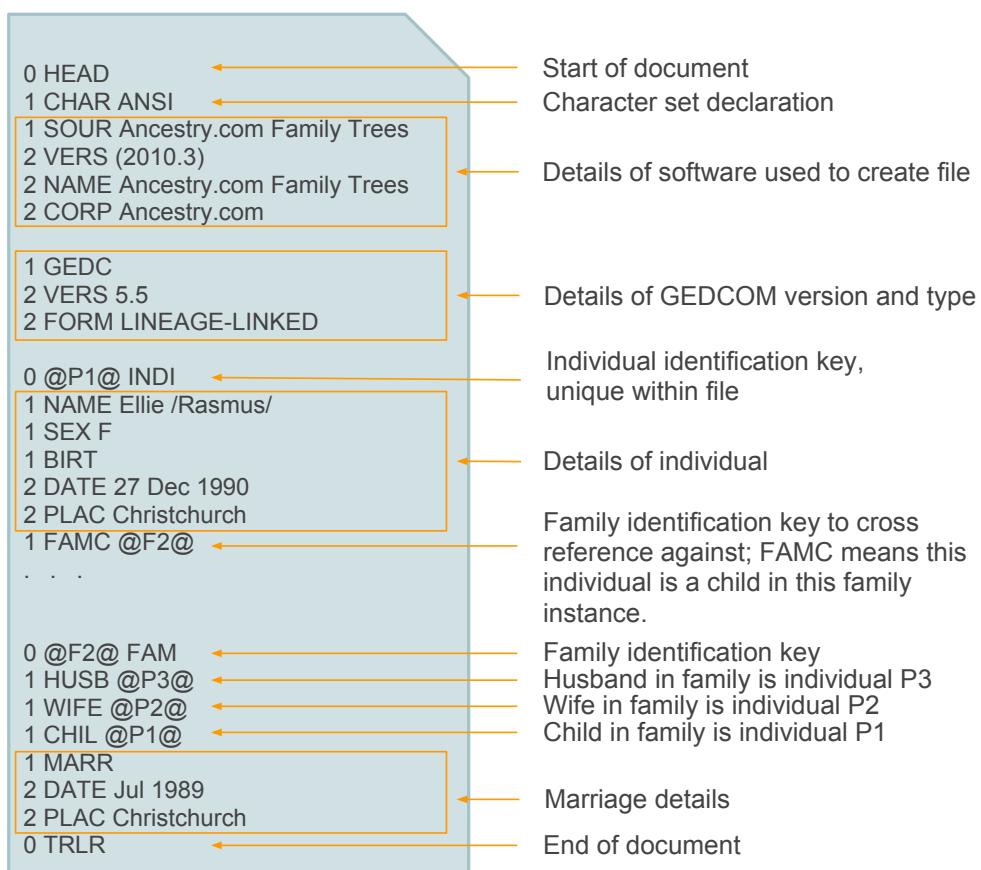


Figure 2.6: An excerpt from a GEDCOM file, illustrating INDIVIDUAL and FAMILY elements.

3

Design and implementation

3.1 Overview

The proposed system can be divided into three parts: the context filter, NeoFace face recognition, and the parsed GEDCOM structure.

3.2 GEDCOM Files

In this project, the gedcom4j library has been used for the parsing of GEDCOM files [22]. This library loads a given GEDCOM file and parses it into a navigable Java object hierarchy. While it is intended for use within Java applications, the gedcom4j library is being used within this C# application through the IKVM.NET implementation of Java for the Microsoft .NET framework [19].

3.3 Face Recognition

3.3.1 Face Grouping with NeoFace

NeoFace is provided with a directory of images. Each one of these images undergoes face detection, where any detected faces are saved in a database (see Table 3.1). Each face is then matched against every other face in the database. This involves loading all other faces as target data, the query face as query data, and verifying the query face's identity with a certain similarity threshold. If the query face matches any target faces, a list of matches and scores will be returned. The face with the highest score, that also does not appear in the same source image as the query face, is considered a match. The query image is then saved under the matching face's face group; all faces in a face group are considered to be of the same individual.

Matching Threshold

The matching or similarity threshold is a parameter, between 0 and 1, used by NeoFace during identity verification. A lower threshold will result in faces of low similarity being considered a match, while a higher threshold will require faces to be very similar for a match to occur. Different face recognition applications will require different matching thresholds. When testing against images taken in a very controlled environment such as passport photographs or police mugshots, the threshold can be much higher so to avoid false positives. However, family albums consist of many photographs taken in uncontrolled environments, resulting in varying pose, illumination and image quality. Five different matching thresholds were tested on our album of 523 faces (see Fig. 3.1). The lowest threshold of 0.6 incorporated many more faces into face groups, which also resulted in a higher error rate. Conversely, a matching

Entity name	Type	Value
ID	int	(Primary Key) Unique identifier for each row
FaceGroupID	int	Links this face to other faces sharing this face group id
ImagePath	string	Location of source image on local system
FaceImagePath	string	Location of extracted face image on local system
FaceScore	float	The probability of this being a face, given by NeoFace
IndiID	string	Once identified, references an individual within a Gedcom file
Gender	string	Face's gender, estimated by AgeAndGender library
Age	int	Face's age, estimated by AgeAndGender library
FeatureData	byte array	Facial feature data extracted by NeoFace
FacePosition	int	(Foreign Key) ID of entry in FacePosition table, locates face within source image
DateFromEXIF	DateTime	The time the photograph was taken, if EXIF metadata is available

Table 3.1: The main database structure, showing information extracted for each detected face.

threshold of 0.75 provided a lower error rate, but only places 114 faces into groups. While we require a certain level of correctness, we also want a minimal number of singletons, or faces in their own face group. When considering the number of faces grouped as well as the number of errors, a threshold of 0.68 is the best choice, and what has been used throughout the rest of the system implementation.

3.3.2 Age and Gender Estimation

Supplementary to NeoFace is the AgeAndGender library, that provides us with estimations of a face's age and gender.¹ When deciding whether a face should be matched to a particular face group, we could use this information to compare the calculated gender of the query image with the gender of the face group. A face with the same gender is more likely to belong in the face group. Estimated ages of identified individuals could also be used in the identification of other people in the photo, by providing a smaller photograph timeframe to acquire candidates from (timeframes are discussed in section 3.4.1). Of the 523 faces in the photo album being tested, gender was accurately predicted for 64.9% of them, with 12.5% of faces having incorrectly predicted gender, and 22.6% of faces having unknown gender. It is worth noting that 14.5% of the faces tested were of children that were either assigned the incorrect gender or whose gender was unknown, showing that this estimation of age and gender struggles more with children's faces. While it is impossible to compute the exact accuracy of the age estimations, as we do not know exactly when each photograph was taken, examining a small number of face groups shows a huge variance in accuracy; while some results were accurate within 2 to 3 years, some results were very inaccurate, with children being assigned ages in their mid-forties, and some elderly faces estimated to be in their thirties (see Fig. 3.2 for examples). Due to the widely varying accuracy of these estimations, it has been decided not to use these measures as context within the system. Possible uses for this information are discussed in section 5.

¹This functionality, and the libraries being used, are still under development by NEC.

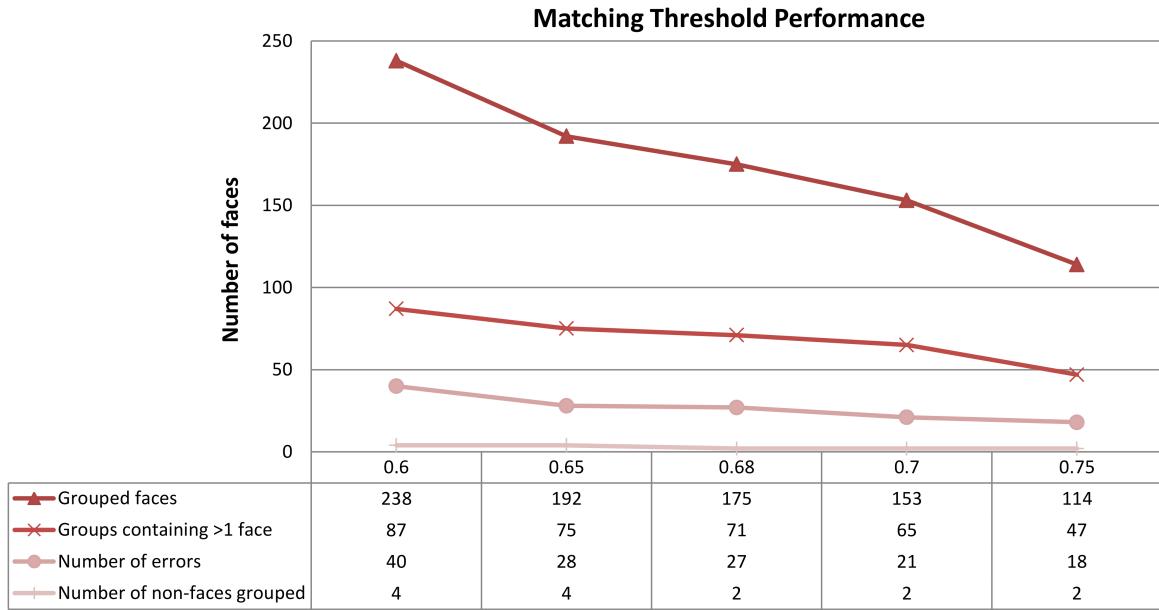


Figure 3.1: The performance of face grouping under differing matching thresholds.

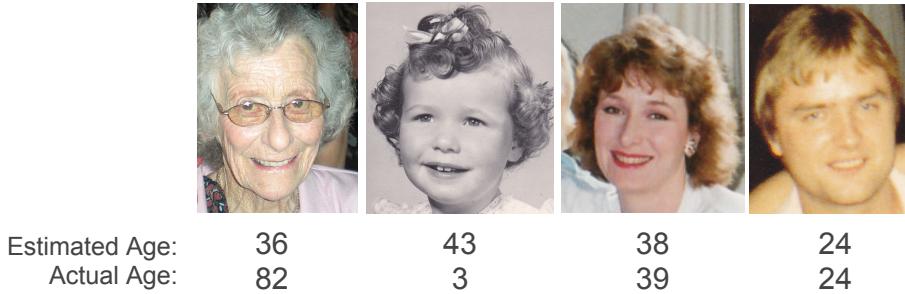


Figure 3.2: Some examples of estimated ages.

3.4 Context Filter

The context filter of this system computes identity suggestions for each face group based on information derived from a GEDCOM file provided alongside the album of photos, as well as contextual information derived from the album itself. These suggestions will be computed based on a series of factors, which can be grouped into five areas; time frames, degree of familial relationship, co-occurrence, recurrence, and physical factors. Figure 3.3 gives a high-level overview of the calculations performed when ranking candidates against a query face.

3.4.1 Time Frames

Historical photos within a family album will most likely have been digitised using a flatbed scanner, meaning that any time metadata within the image will refer to the time of digitisation, not its original creation date. Because of this, we often do not have any accurate data regarding a photo's creation time, and another method of estimating a photo's possible time

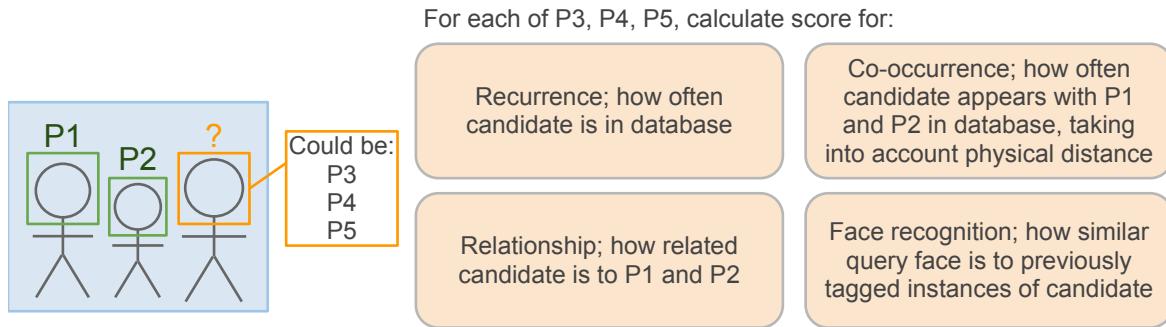


Figure 3.3: An overview of the candidate scoring process.

of creation is necessary. In this system, a time frame is a pair of dates denoting a period of time within which a photograph must have been taken or an individual must have been alive, according to information currently known about identified individuals and the photographs they appear in.

Photograph Time Frames

The time frame of a photograph represents the possible period of time within which the photo must have been taken. This is essentially the period of time where all individuals present in the photograph were alive. The more individuals identified in the photograph, the narrower and more accurate the time frame will be. For example, consider a photograph containing a number of individuals, but only 3 currently identified. Person 1 was born in 1900 and died in 1980, person 2 was born in 1920 and died in 1990, and person 3 was born in 1950 and is still living. From this, we know that the photo must have been taken between 1950 and 1980, as this is the time frame within which all identified individuals were living. If a fourth person is then identified who was born in 1975, we know that this photograph must have been taken between 1975 and 1980, narrowing the time frame significantly. When identifying remaining people, we can use this information to disregard any candidates that were not living during this time.

Face Group Time Frames

The time frame of a face group consists of the earliest and latest possible dates that an individual could have lived, based on the photographs they appear in and who they appear with. This is achieved by accumulating the time frames of all photographs the individual appears in. By taking the earliest start date and latest end date of time frames of photographs that the individual is present in, we can construct a time frame for each individual.

Candidate lists

When a suggestion is to be made as to the identity of an unknown face, a list of possible candidates is generated. This is initially all of the individuals in the GEDCOM file that were alive during the time frame of the photo. These candidates are then scored and ranked; this will be covered in depth in later sections.

The manner in which time frames are calculated means that they are more susceptible to type 1 errors; it is likely that some unsuitable candidates will remain in the candidate list as a result of the photo time frame being too broad. This could either be because few individuals have already been identified in the photograph, the appearance of an individual who lived for a long period of time, or a group of people that all lived around the same time, meaning that the time frame cannot be narrowed down.

In the case of no individuals having been previously identified in a photograph, the initial candidate list will be very large, consisting of all individuals from the GEDCOM file alive after the invention of photography.²

3.4.2 Degree of Familial Relationship

From the GEDCOM file, we can calculate the coefficient of relationship between each individual in the family tree. This value indicates the theoretical proportion of genes two individuals share.

Calculating the Coefficient of Relationship

To find the coefficient of relationship between two individuals a and b , their lowest common ancestor (LCA) must first be identified; this is done by retrieving a list of all ancestors of each individual and iterating through the lists, finding the first instance of an individual appearing in both ancestor lists.

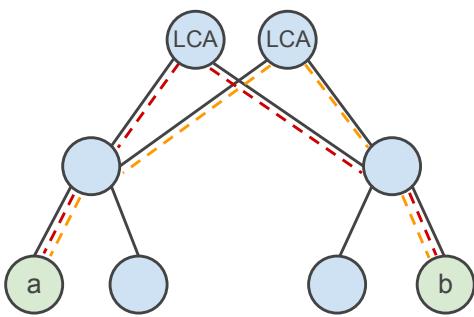


Figure 3.4: Nodes a and b are first cousins, and are separated by 4 paths through each grandparent.

This means that these two individuals theoretically share 12.5% of their genes with one another. The relationship between parent a and child b is calculated by $(1/2)^1$. Summing of multiple paths is not required, as a is b 's direct ancestor. This results in the child theoretically sharing 50% their genes with their parent.

There are some situations where the standard coefficient of relationship is not a fair representation of the relationship between two people. For example, spouses would (assuming no inbreeding) share no common ancestor, resulting in a relationship measure of 0. However,

The number of paths separating the two individuals can be found by adding the distance of a from the LCA to the distance of b from the LCA (see Fig. 3.4).

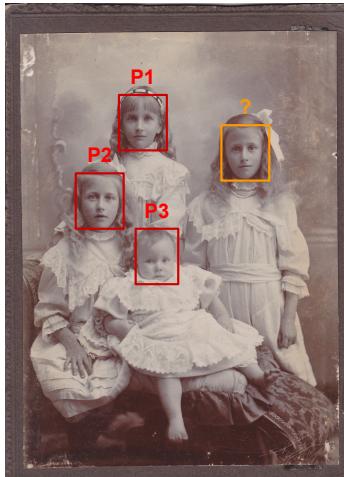
The coefficient of relationship can then be calculated by $R_{ab} = \sum(1/2)^n$, where n is the number of paths separating individuals a and b . Summing of all possible paths is required for relationships where individuals share more than one lowest common ancestor. For the example illustrated in Fig. 3.4, individuals a and b are separated by 4 paths through each grandparent. Their degree of relationship can be calculated by $(1/2)^4 + (1/2)^4$, resulting in a value of 0.125.

²The first permanent photograph was created in 1826 by Joseph Nicéphore Niépce; January 1st 1826 has been used to represent the invention of photography

a spousal relationship is arguably one of the most important in terms of co-occurrence; in the album of 173 family photographs being tested, spousal relationships appear at a similar rate to parent/child and sibling relationships (see Fig. 3.2). For this reason, it has been decided to score spousal relationships with the same value as parent/child and sibling relationships.

Relationship	Frequency
Sibling	70
Spouse	65
Parent/child	61
Grandparent/child	31
Cousin	31

Table 3.2: The rate of occurrence of different relationships within the album.



For each candidate i , average relatedness =

$$\frac{R(P1, i) + R(P2, i) + R(P3, i)}{3}$$

Figure 3.5: Calculating a candidate's average relatedness to the identified individuals in a photograph.

Using the Coefficient of Relationship

Within a photograph containing more than one individual, a cumulative relationship coefficient for each candidate can be calculated by adding together the relationships between each candidate identity and each other individual in the photograph. For example, a photograph featuring individuals $P1$, $P2$, $P3$ and an unknown individual x (see Fig. 3.5). For each individual in the candidate list for x , the coefficient of relationship is calculated for each person already identified in the photograph. Finding the average of these relationships provides the relatedness of each candidate to the group of identified people in the photograph. A candidate with a higher relatedness to the group is more likely to appear in the photograph.

3.4.3 Co-occurrence

In this system, co-occurrence is a relative measure of how often an individual appears alongside every other individual throughout a set of photographs.

Number of people in photograph

The co-occurrence of two individuals within a photograph is of more importance if there are less other people in the photograph. For example, a couple appearing alongside each other in a wedding photograph is more significant than two people appearing in a large group photograph at a family reunion.

Physical Distance

The physical distance between two individuals when they appear in photographs can be an important indication of their relationship. If two people routinely appear close together in many photographs, there is a higher chance that they are closely related or spouses. The euclidean distance between two faces p_i and p_j must first be normalised to ensure a meaningful measurement. This can be done by taking into account the average face width within the photograph in pixels and a standard width of a face in centimetres.

Also taken into account is the average distance between all the faces in the photograph (see Fig. 3.6). This means that co-occurrence is more significant if two people appear close together in a photo where most people are far apart.

3.4.4 Recurrence

Recurrence is a simple measure of how often an individual appears throughout the album compared to the average rate of appearance. Individuals that appear more frequently within an album are more likely to appear again.

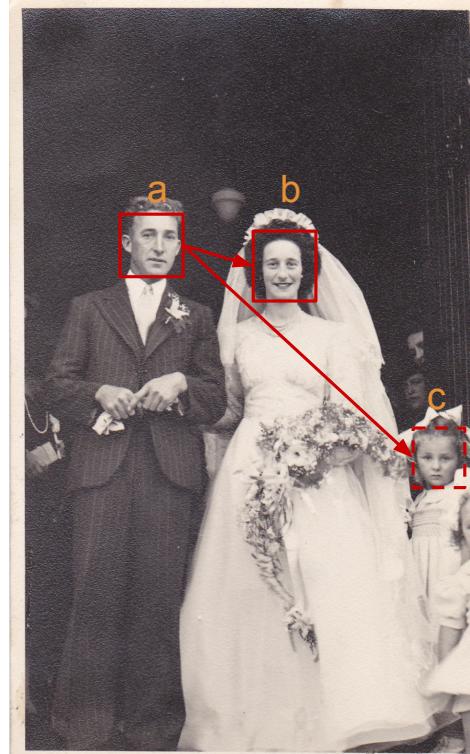


Figure 3.6: When using physical distance as a measure of context, we take into account the distance between all of the faces in the photo.

3.5 System Structure

The system itself is implemented in four files: GedcomTree.cs, NeoFace.cs, ContextFilter.cs, and Form.cs (See Fig. 3.8).

The initial input provided to the system will be in the form of a GEDCOM file and an album of untagged family photographs (see Fig. 3.7 for an overview of operations performed on initial startup). While we have this data, we cannot use any of the data from the GEDCOM file in a meaningful way as we have not yet identified any individuals. The first action of the system will be to detect and extract faces from the source images. These faces will be grouped together by similarity of appearance. Some photos may contain very small faces that NeoFace cannot extract meaningful feature data from. These faces are detected and extracted by the supplementary AgeAndGender library and are stored in the database, but cannot be used for face recognition queries.

From the data now in the database we can calculate co-occurrence rates. This measures the frequency with which each face group, currently an unidentified individual, appears alongside each other unidentified individual in the photo album.

We can also compute various attributes of each face group that are not dependent on any data from the GEDCOM file. Each face group is assigned an arbitrary identification number, which will later be mapped to an INDI identifier within the GEDCOM file. The recurrence measure of a face group can be calculated by dividing the number of faces within the face group by the total number of faces within all photographs. This can later be compared to the average recurrence of faces in the album.

3.6 Composing the final score

To ensure effective use of these various scores, we need to devise an appropriate distribution for their summation. While testing accuracy using face recognition alone, we find that the average likeness score given to top ranking candidates is 40.28. No two individuals can have a relationship score higher than 50, and the average relatedness of a candidate to a group of individuals in an image is 9. The average co-occurrence score is 15.5, and the average recurrence rate is 19.2.

Initial tests show that changing the weighting of the relationship score alone does not cause any major change in performance (see Fig. 3.9). As face recognition provides a high level of accuracy by itself, we do not want these context scores to hold too much weight relative to the face recognition scores.

Various combinations of face recognition and context scores have been tested. In Fig. 3.10 we see the average accuracy of identity ranking using 11 different score compositions, with Context referring to the combination of co-occurrence, recurrence, physical distance and

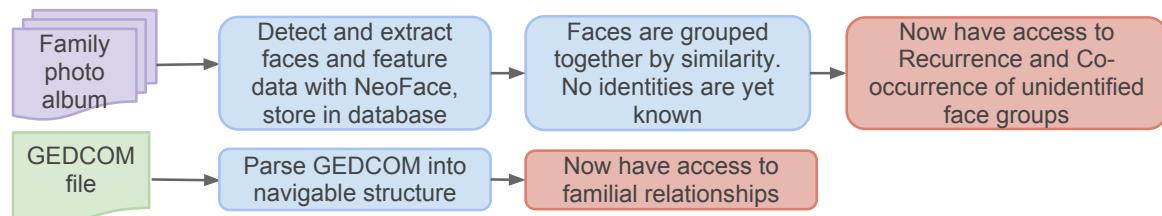


Figure 3.7: Tasks to be performed on initial startup.

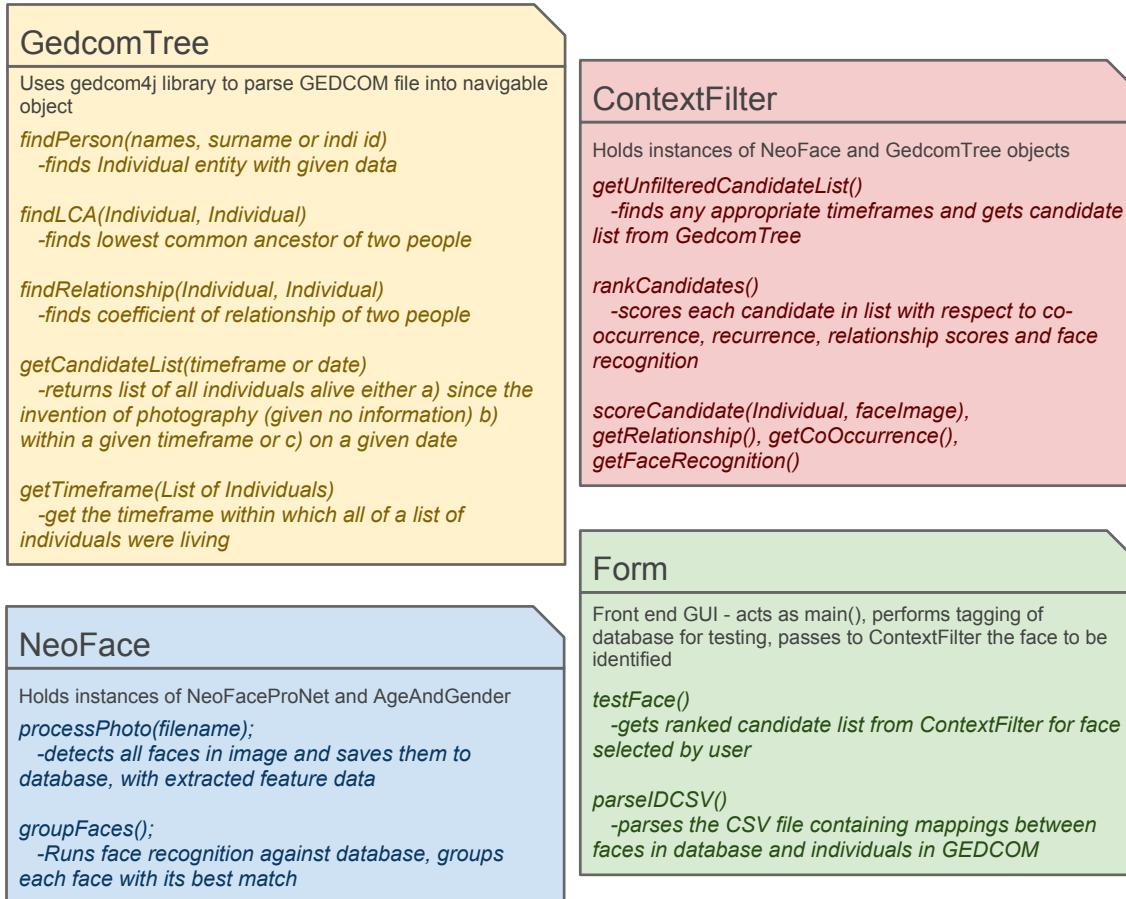


Figure 3.8: Outline of classes the system is composed of.

relationship measures. In this test, we want values closer to 1, as a ranking of 1 means that the correct identity is deemed the most likely to belong to the query face. Giving the face recognition score double the weight of the context score provides the lowest average rank, with the worst performances coming from the use of face recognition alone, and giving the context score 3 times the weight of the face recognition score.

From these results we settle on the use of the context score alongside 2 times the face recognition score. This means the final identity ranking formula can be expressed as:

$$FinalScore = \left[\begin{array}{l} PhysicalDistanceScore \\ +CoOccurrenceScore \\ +RelationshipScore \\ +RecurrenceScore \end{array} \right] + 2 \times FaceRecognitionScore$$

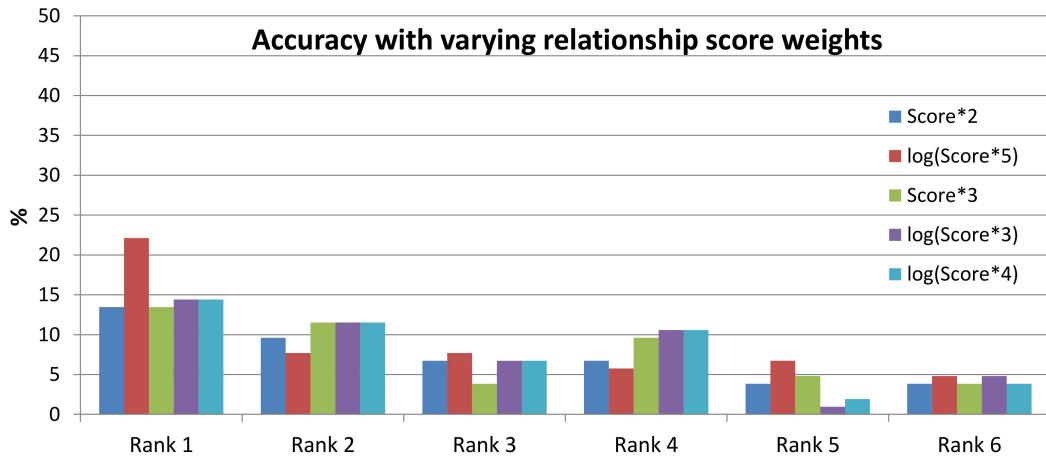


Figure 3.9: The distribution of identity ranks when using different relationship score weights.

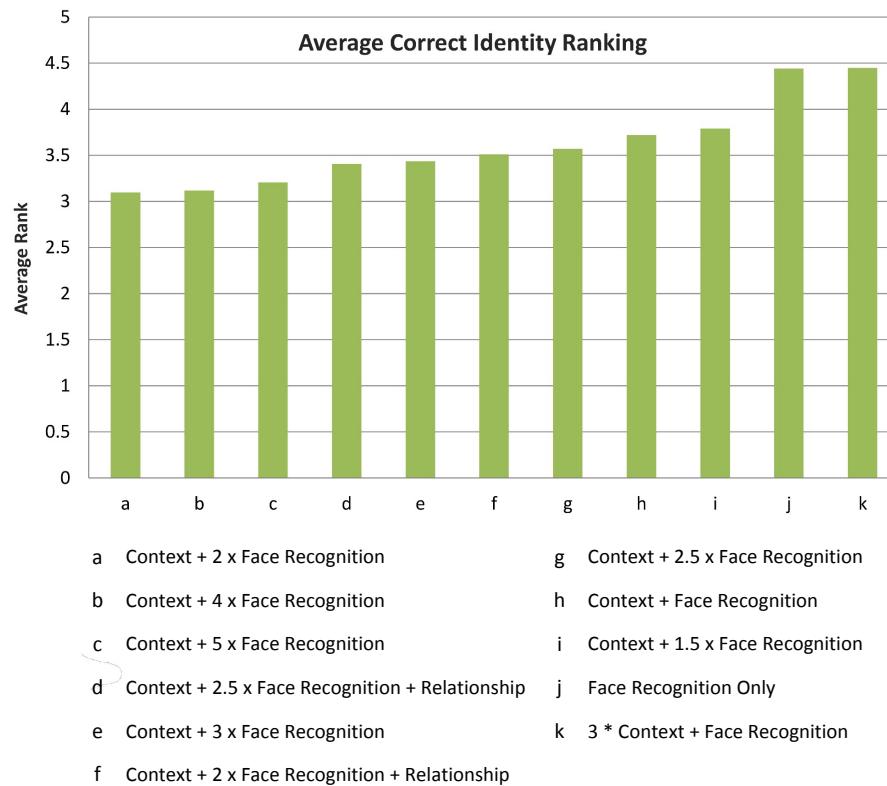


Figure 3.10: The average accuracy of identity predictions using various score compositions.

4 Results

4.1 Experimental Design

Testing has been performed on a database of 523 faces, 349 of which have been used as query faces. Not all of these 523 faces in the database were suitable for use as query faces; some faces extracted from the photo album are of individuals that do not appear in the GEDCOM file or are of unknown identity. Also, we cannot use as query faces any faces detected by the AgeAndGender library, as they are not accompanied by feature data that can be used for facial recognition. These faces are retained in the database as they can assist with context measures, and help to maintain the structure of an actual family photo album.

The database starts with all known faces being tagged as individuals in the GEDCOM file. The GEDCOM file used for testing consists of 948 individuals, 60 of which are present within photos in the album. For each of the 349 faces tested against the database, the face being tested is untagged (or set as being of unknown identity), while the rest of the database remains tagged. The system then predicts the query face's identity, the face is retagged, and the process is repeated with the next face. The main output being evaluated is the position in the final candidate list of the query face's actual identity. The best outcome is the query face's identity being in rank 1.

In this evaluation, a ranking of 50 has been used as a threshold; any face with its correct identity ranked worse than 50 will not appear on graphs or in data analysis. Of these 349 faces, there were 19 outliers that resulted in an identity ranking lower than 50.

The performance of NeoFace's face recognition alone is used as the baseline for testing the accuracy of identity prediction. The performance of the proposed system is shown through the use of context in scoring alongside the face recognition score.¹

Testing has been performed on a machine running 32-bit Windows 7 SP 1 with 4 GB RAM. Face recognition is performed with the NeoFace version 3.1 runtime environment, and system development is in C# using Microsoft Visual Studio 2010.

4.2 Overall Accuracy

Figure 4.1 shows the performance of the system in terms of accuracy with respect to the top 50 ranked candidate identities for each query face. A hit list of size 1 means that the correct identity is ranked as the most likely identity of the query face, while a hit list of size 50 requires the correct identity of the query face to be ranked in the top 50 identities. The two data series shown correspond to the accuracy of using face recognition alone to score and rank candidates, and the accuracy when using context measures as well as face recognition in scoring. We find that scoring with context measures compared to the use of face recognition alone achieves a

¹Throughout these results, the performance of face recognition alone is shown in red and the proposed system's performance is shown in blue

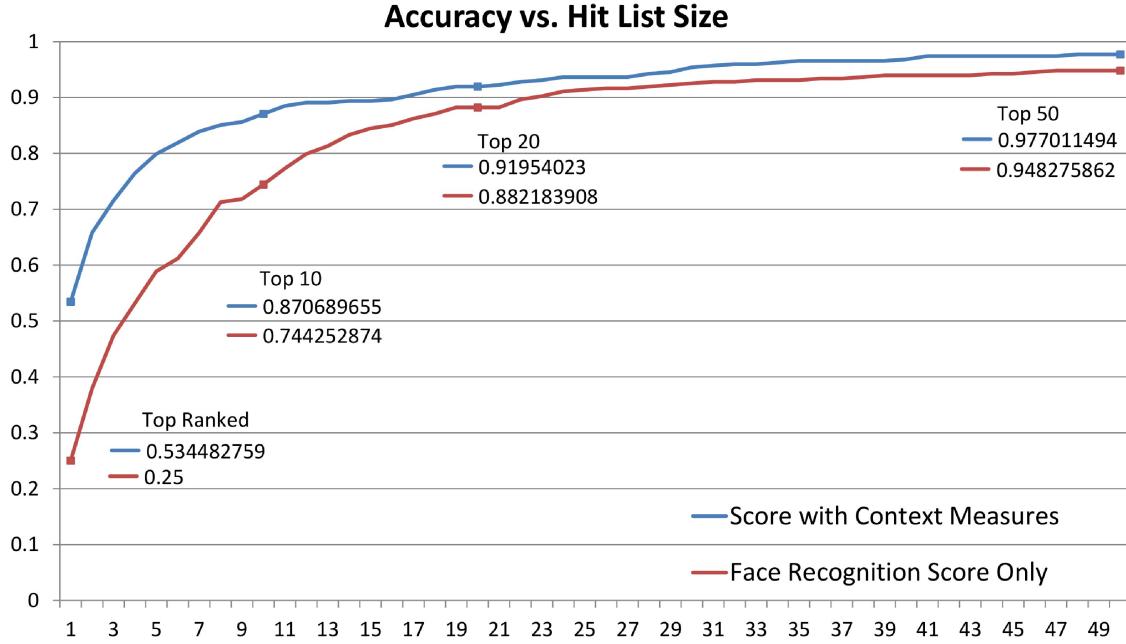


Figure 4.1: Accuracy of identity prediction at different hit list sizes.

26% improvement in accuracy at hit list size 1, a 21% improvement at hit list size 5 and a 12% improvement at hit list size 10. While the performance of face recognition alone tends towards a similar accuracy as scoring with context as hit list size increases, we are more interested in the system’s performance using a hit list of only 5 or 10 identities, as this is more useful in terms of realistic face tagging interfaces. Overall we find a significant improvement in mean ranking with the use of context measures (paired T-Test, $T = 8.9179, p < 0.001$; Wilcoxon $z = 9.226272, p < 0.0000001$).

4.3 Accuracy vs. Identity Recurrence

Figure 4.2 shows the accuracy of identity prediction with respect to the frequency of the query face’s identity in the database. As we can see by the red trendline, there is no significant improvement in the performance of using face recognition alone when there are many instances of the query face’s identity (Kruskal-Wallis Test corrected for ties, $H = 7.335848, df = 6, N_1 = 53, N_2 = 58, N_3 = 41, N_4 = 35, N_5 = 25, N_6 = 17, N_7 = 100, p = 0.290902$). However we can see a significant relation between identity frequency and accuracy when using context measures in scoring (Kruskal-Wallis Test corrected for ties, $H = 88.460953, df = 6, N_1 = 53, N_2 = 58, N_3 = 41, N_4 = 35, N_5 = 25, N_6 = 17, N_7 = 100, p < 0.0000001$).

4.4 Accuracy vs. Candidate Co-occurrence

Figure 4.3 shows the accuracy of identity prediction with respect to the correct identity’s co-occurrence score, or how often the correct identity appears throughout the album alongside the identified individuals in the photo. The red trendline shows no significant change in accuracy as identity co-occurrence increases when scoring using face recognition alone (Kruskal-Wallis Test corrected for ties, $H = 10.025022, df = 6, N_1 = 49, N_2 = 106, N_3 = 65, N_4 = 49, N_5 =$

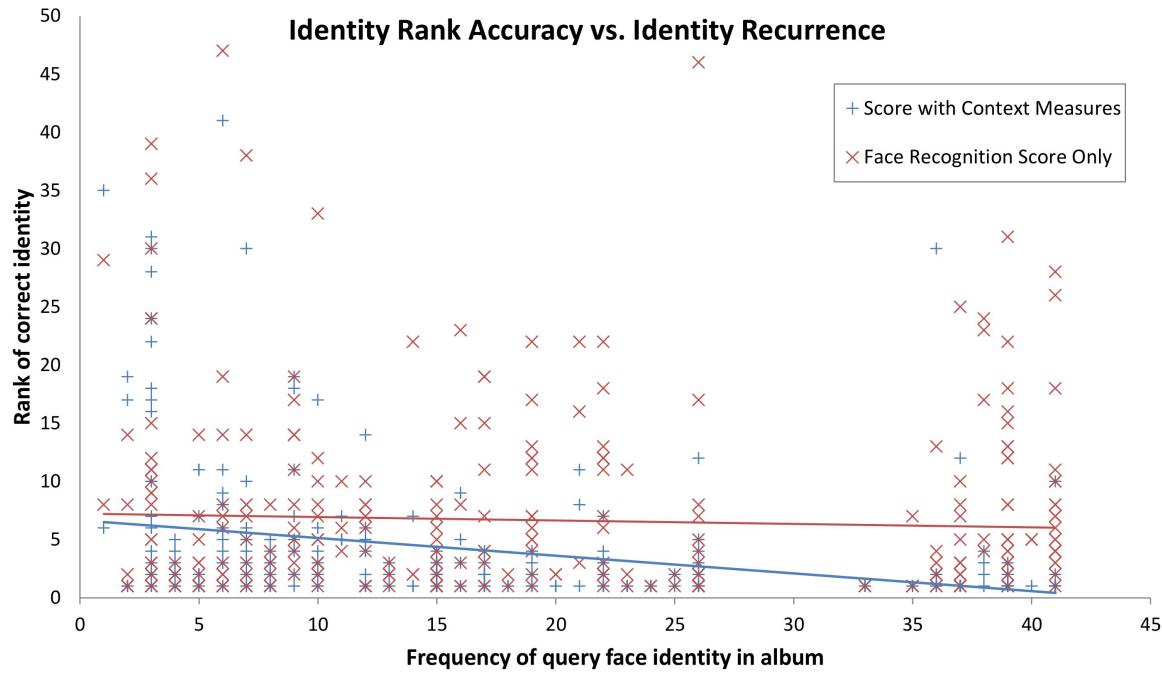


Figure 4.2: Accuracy of identity prediction with respect to rate of identity recurrence in album.

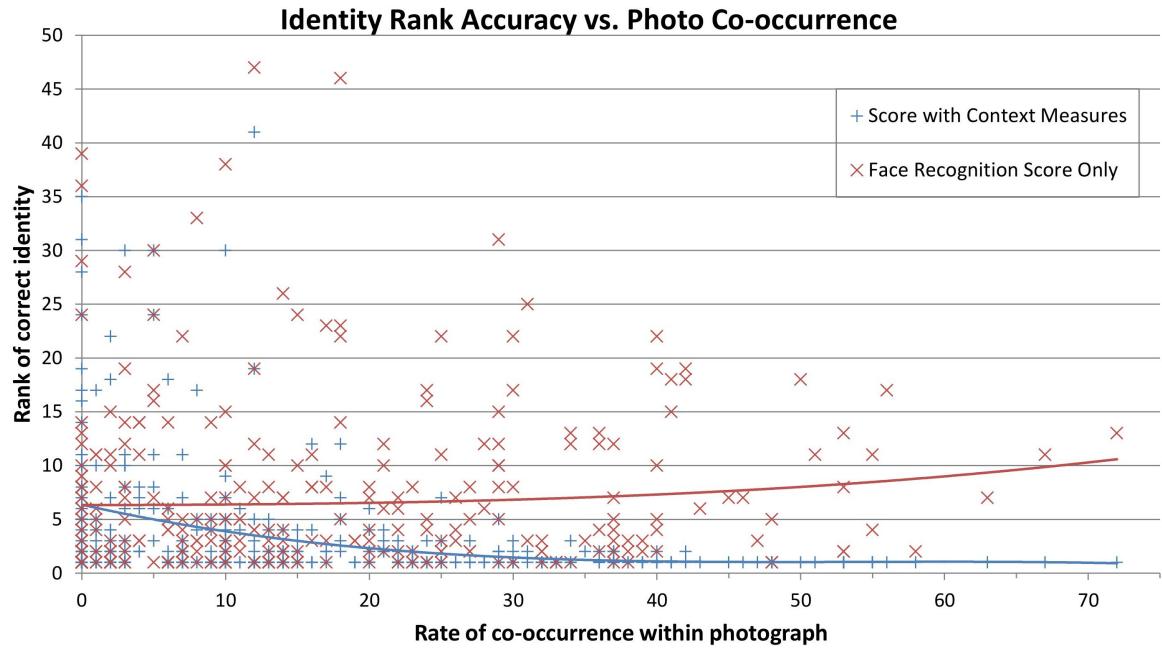


Figure 4.3: Accuracy of identity prediction with respect to the correct identity's co-occurrence score.

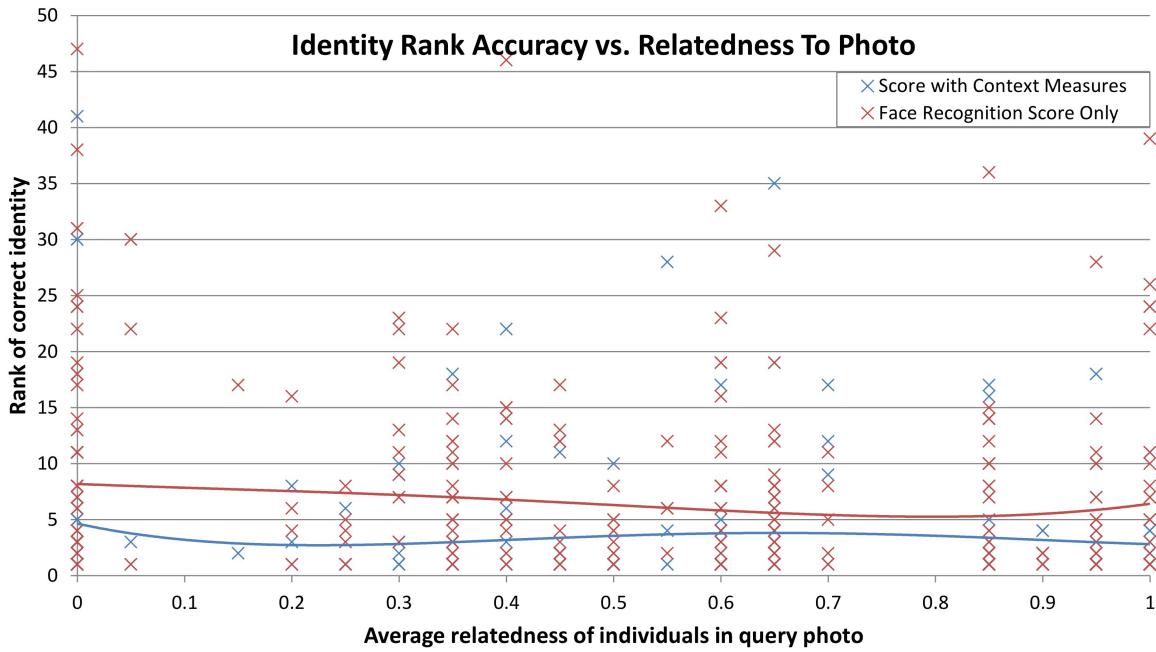


Figure 4.4: Accuracy of identity prediction with respect to the relatedness of the query identity to the rest of the individuals in the photo.

$37, N_6 = 12, N_7 = 12, p = 0.123602$). The blue trendline, however, shows a definite increase in accuracy as the rate of co-occurrence increases (Kruskal-Wallis Test corrected for ties, $H = 48.149176, df = 6, N_1 = 49, N_2 = 106, N_3 = 65, N_4 = 49, N_5 = 37, N_6 = 12, N_7 = 12, p < 0.0000001$).

4.5 Accuracy vs. Relatedness To Photo

Figure 4.4 illustrates the relationship of accuracy to how related the query face identity is to the rest of the photo.

Neither of the trendlines indicate much of a relationship between accuracy and relatedness. Analysis shows relatedness having more of an impact on the accuracy of face recognition scoring alone (Kruskal-Wallis Test corrected for ties, $H = 29.184748, df = 16, p = 0.022720$) compared to scoring with context measures (Kruskal-Wallis Test corrected for ties, $H = 19.258095, df = 16, p = 0.255507$). This is clearly nonsensical, as scoring solely with face recognition does not involve any relationship measures at all. We dismiss this as coincidental, and accept that the relatedness of the query individual to the individuals in the query photo does not impact accuracy.

4.6 Accuracy vs. Candidate List Size

Figure 4.5 shows the identity prediction accuracy as the number of candidates being considered changes. The average candidate list size was 427.

The blue trendline shows a decrease in performance as the size of the candidate list increases when using context measures in scoring (Kruskal-Wallis Test corrected for ties, $H = 23.601916, df = 5, N_1 = 3, N_2 = 9, N_3 = 145, N_4 = 106, N_5 = 48, N_6 = 18, p = 0.000259$).

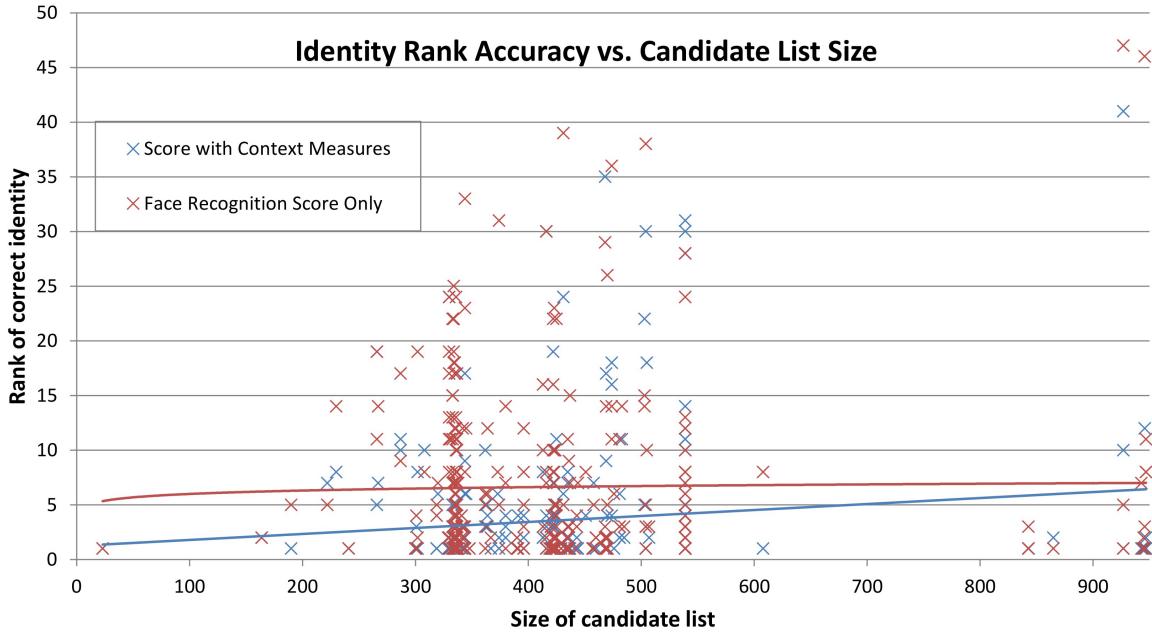


Figure 4.5: Accuracy of identity prediction with respect to the number of candidates considered.

This is to be expected; the less candidates there are to consider, the more likely that the applied context measures will successfully apply to only a small proportion of the candidate list, and the smaller the maximum possible rank. We see that candidate list size has an insignificant impact on scoring with face recognition alone (Kruskal-Wallis Test corrected for ties, $H = 8.017898, df = 5, N_1 = 3, N_2 = 9, N_3 = 145, N_4 = 106, N_5 = 48, N_6 = 18, p = 0.155252$).

4.7 Accuracy of Young Face Identification

Dealing with children's faces is a problem that face recognition traditionally struggles with. Figure 4.6 shows the accuracy of identity prediction when dealing with faces under the age of 5, 10 and 15 years of age. In this dataset, 54 faces were under the age of 5, 84 under the age of 10 and 106 under the age of 15. When considering only the top ranking identities (or hit list size 1), scoring with context measures provides a large increase in accuracy, with a 38% improvement in faces under 5, a 42% improvement in faces under 10, and a 40% improvement in faces under 15. This is more significant than the 26% improvement seen across the whole dataset with the use of context. When considering the top 5 ranked identities we see improvements of 20%, 34% and 36% at ages under 5, under 10 and under 15 respectively.

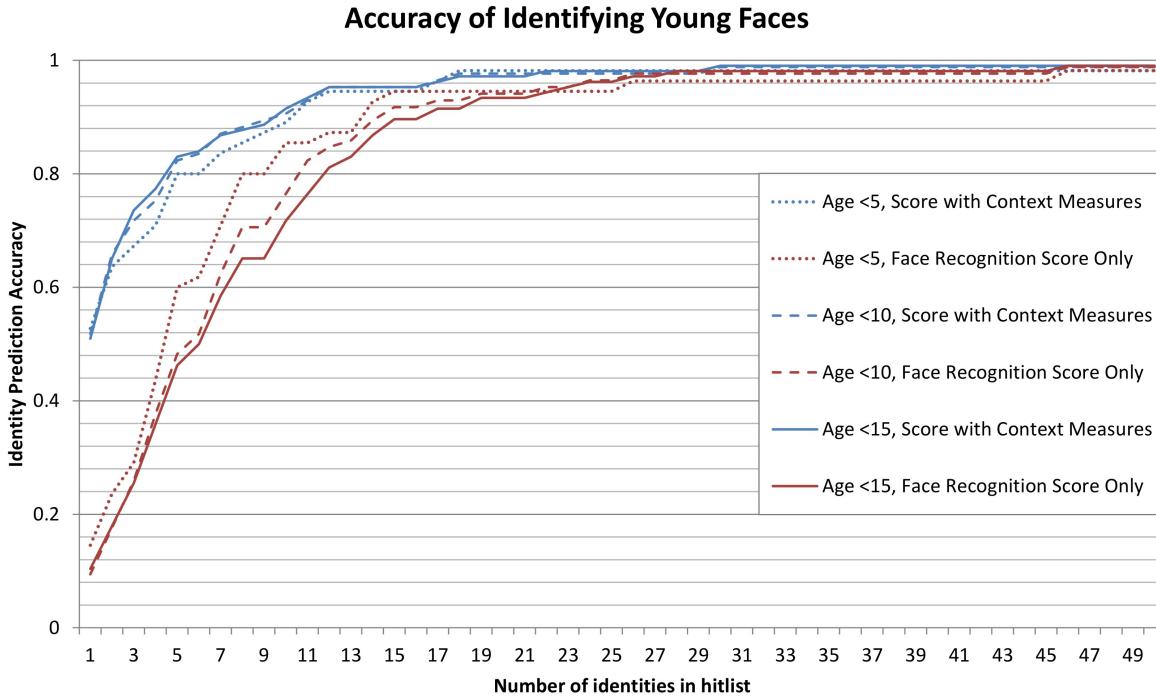


Figure 4.6: Accuracy of identity prediction when dealing with young faces.

4.8 Discussion

We can see from these results that the use of context alongside face recognition provides a clear improvement in identity prediction accuracy. Co-occurrence and recurrence are the most important measures contributing to accuracy. We do not find any strong relationship between accuracy and the relatedness of query faces to the identified individuals alongside them in the photo. Smaller candidate lists tend to provide higher accuracy.

We can also conclude that the use of context effectively overcomes the difficulties that face recognition typically struggles with when identifying young faces.

A limitation of this system is lessened performance when there are few or no instances of the individual previously identified in the database. This, however, is somewhat inevitable, and is a limitation inherently present in many similar systems [47, 11, 35].

The dataset It is worth noting how very challenging this dataset is. The oldest photos in the album date back to around 1908. Faces as small as 32x32 pixels have been correctly identified, and many of the faces are in severe poses, under poor illumination, and suffer from partial occlusion. Figure 4.7 gives an example of some of the faces in the photo album.

Outliers Of the 349 faces used in testing, 19 returned a correct identity ranking of 50 or worse. All 19 of these faces are of individuals of distant relation to the majority of the album. Seven of these faces are of individuals that only appear in the album once, providing no chance for NeoFace to successfully match the face to someone of the same identity. Four of the individuals only have one other face instance in the database, and four of the faces are

alone in their source photo, meaning very little additional context can be drawn from the photo itself.

A note on comparative evaluations While it would be desirable to compare the performance of this system to that of previously published work, this is near to impossible without access to either a) the published systems or b) the set of photographs used for evaluating them, and an associated family tree. This work is reliant on the make up of various relationships within photo albums, as well as the structure, size and time frame of the GEDCOM file accompanying the album and the quality and age of the photos being analysed. For these reasons, results would undoubtedly vary greatly between albums and families, and any meaningful comparison in performance is not possible.

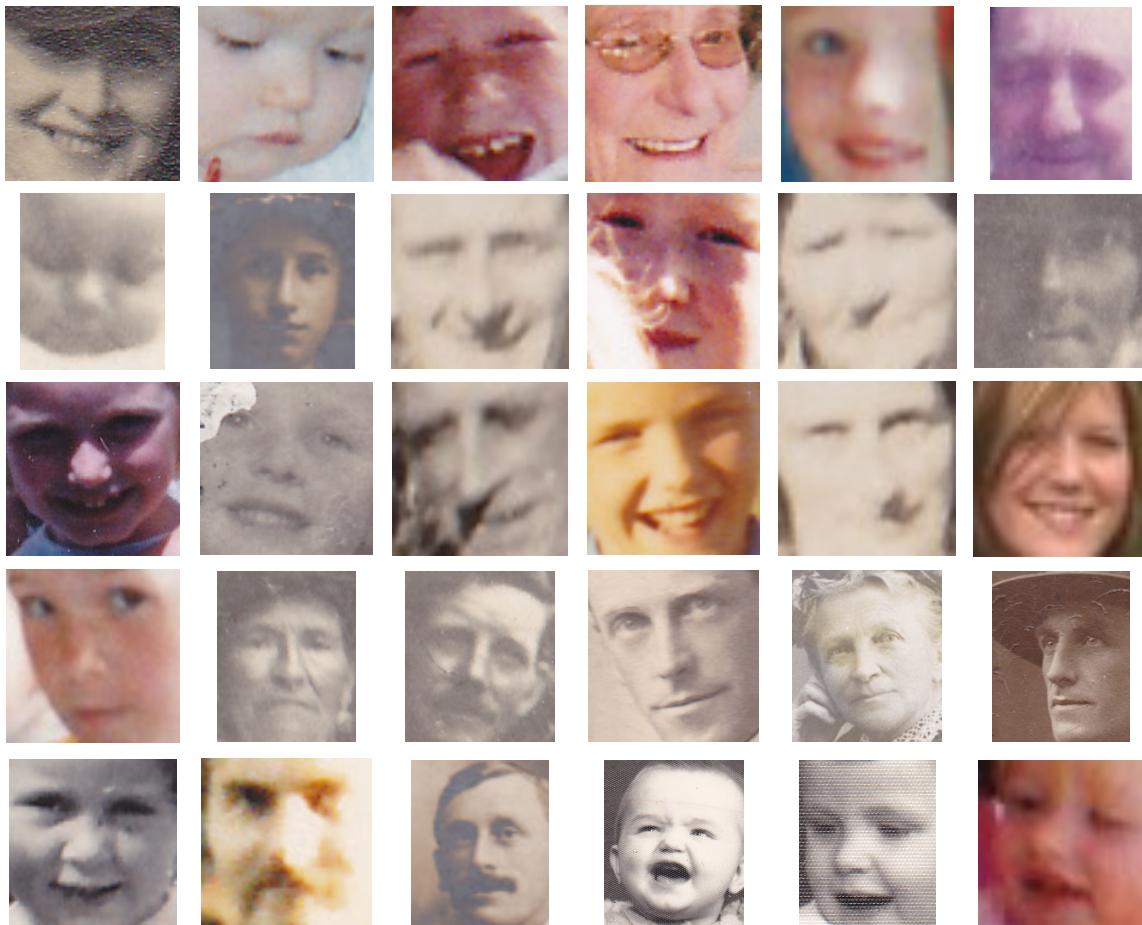


Figure 4.7: An example of some of the faces in the dataset being used for testing

5

Conclusion and Future Work

A novel approach to identity prediction within a family photo album has been presented, making use of relationship measures derived from genealogical information located within a GEDCOM file as well as contextual information about identity co-occurrence, recurrence and relative physical distance derived from the photos within the album. The proposed system, which uses these measures alongside NEC's NeoFace face recognition engine, has provided a 26% improvement in accuracy over the use of face recognition alone, when identifying 348 faces against a database of 523 faces from a challenging dataset of 173 family photographs. Measures of identity co-occurrence and recurrence within the photo album are shown to be an effective addition to identity prediction. The removal of ineligible individuals from a unknown face's list of candidate identities through the construction of photo timeframes is also seen to be an effective way of improving identity prediction accuracy. We see accuracy improvements of up to 42% when identifying the faces of children and young people, something that even the most advanced face recognition technology currently available struggles with.

Future work

This area of research presents a lot of potential future work. One aspect that could be developed further is research about the structure of family photo albums. Areas such as recurring patterns, rates of occurrence, and the most important relationships could be looked into further in order to pinpoint the aspects of family albums that are most beneficial for identity prediction. A lot of analysis has gone into tailoring the composition of the candidate score based on the family album being used for testing, making it difficult to know whether or not these values will need to change in order to create a successful generalised system. Further research into family album structure, alongside testing with multiple datasets, will be helpful for developing a system that can operate successfully over different photo albums.

A current limitation of the system is that it only attempts to identify one unknown query face at a time. It would be beneficial to allow for estimations to be made for a pair or multiple unknown faces simultaneously; this way, the ranked candidate lists of both unknown faces could influence each other. For example, if there are two unknown faces appearing alongside each other in an image, and there are two highly ranked identities in the candidate lists that are married to each other, it is more likely that these unknown faces belong to this couple, even if the identities are not the highest ranked in each of their respective candidate lists.

It would be very helpful in the future to incorporate estimates of the age and gender of faces. As these measures are often not robust when applied to such a varying dataset, measures such as the face image resolution, the angle of face pose, and face detection certainty could be taken into account when deciding how much weight to give the estimations.

When calculating the relatedness of a candidate identity, measuring the relatedness within the group in the photo could provide additional information. If all of the individuals previously

identified in a photograph are closely related, the coefficient of relationship will be of higher importance when ranking the candidate list. If a photograph contains individuals that are not closely related or are not related at all, the coefficient of relationship calculated for each candidate is not going to be of such importance. For example, the coefficient of relationship is not going to hold great relevance when identifying individuals in a school class photograph, as the people in the photograph are bound by a social, rather than familial, relationship.

Another measure of context that could be helpful for predicting identities is that of relative face size. Some relationships can be identified by comparing the size of an unknown face to the other faces in the photo, as well as its position relative to those faces. For example, when a parent holds a young child, the child's face is smaller and appears below the parent's face. Similarly, a couple standing next to each other will often be slightly different heights and of similar face size. Taking this into account could help to identify relationships within photos.

If such a system was being developed for release rather than research purposes, it would be necessary to allow the user to tag individuals not present in the GEDCOM file. This could be done either by providing functionality to add new individuals to the family tree, or by maintaining a secondary set of individuals alongside the GEDCOM file that are not necessarily family members, but appear in the photo album.

Shao et al. use a series of Gabor filters to extract genetic-invariant features, in order to verify kinship using the UB KinFace database. This consists of images of children and parents at varying ages [30]. While this is out of the scope of this paper, it is an area of future interest that could be incorporated into the facial recognition aspect of the system to assist in identifying related individuals.

Bibliography

- [1] faint: The Face Annotation Interface, June 2012. <http://faint.sourceforge.net/>.
- [2] Family Tree Maker by Ancestry.com, September 2012. <http://www.familytreemaker.com/>.
- [3] Gramps: Genealogical Research Software, September 2012. <http://gramps-project.org/>.
- [4] How face recognition works - MyHeritage.com, October 2012. <http://www.myheritage.com/facial-recognition-technology>.
- [5] MyHeritage, June 2012. <http://www.myheritage.com/>.
- [6] Picasa and Picasa Web Albums, October 2012. <http://support.google.com/picasa/>.
- [7] S. Aly, A. Sagheer, N. Tsuruta, and R.-i. Taniguchi. Face recognition across illumination. *Artificial Life and Robotics*, 12(1):33–37, 2008.
- [8] B. Ana Paula, L. O, O. Camillo Jorge Santos, A. Arnaldo de Albuquerque, and Jo. Face Recognition Aiding Historical Photographs Indexing Using a Two-Stage Training Scheme and an Enhanced Distance Measure, 2008.
- [9] D. Anguelov, K. Lee, S. Gokturk, and B. Sumengen. Contextual identity recognition in personal photo albums. In *Computer Vision and Pattern Recognition, 2007. CVPR '07. IEEE Conference on*, pages 1 –7, june 2007.
- [10] D. J. Beymer. Face recognition under varying pose. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pages 756–761, 1994.
- [11] M. Brenner and E. Izquierdo. Graph-based recognition in photo collections using social semantics. In *Proceedings of the 2011 ACM workshop on Social and behavioural networked media access*, SBNMA '11, pages 47–52, New York, NY, USA, 2011. ACM.
- [12] W. Cao, R. Che, and D. Ye. An illumination-independent edge detection and fuzzy enhancement algorithm based on wavelet transform for non-uniform weak illumination images. *Pattern Recognition Letters*, 29(3):192–199, 2008.
- [13] R. F. Carvalho, S. Chapman, and F. Ciravegna. Attributing semantics to personal photographs. *Multimedia Tools Appl.*, 42(1):73–96, Mar. 2009.
- [14] K. Chandramouli and E. Izquierdo. Semantic structuring and retrieval of event chapters in social photo collections. In *Proceedings of the international conference on Multimedia information retrieval*, MIR '10, pages 507–516, New York, NY, USA, 2010. ACM.

- [15] T. Chi-Yao and C. Ming-Syan. Photo identity tag suggestion using only social network context on large-scale web services. In *Multimedia and Expo (ICME), 2011 IEEE International Conference on*, pages 1–4, 2011.
- [16] R. Dale-Jones and T. Tjahjadi. A study and modification of the local histogram equalization algorithm. *Pattern Recognition*, 26(9):1373–1381, 1993.
- [17] L. Ding and A. Yilmaz. Inferring social relations from visual concepts. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 699 –706, nov. 2011.
- [18] N. Elahi, R. Karlsen, and W. Younas. Image annotation by leveraging the social context. In *Proceedings of the 5th International Conference on Ubiquitous Information Management and Communication*, ICUIMC ’11, pages 40:1–40:8, New York, NY, USA, 2011. ACM.
- [19] J. Frijters. IKVM.NET, July 2012. <http://www.ikvm.net/>.
- [20] A. S. Georghiades, P. N. Belhumeur, and D. J. Kriegman. From few to many: illumination cone models for face recognition under variable lighting and pose. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(6):643–660, 2001.
- [21] Y. Z. Goh, A. B. J. Teoh, and M. K. O. Goh. Wavelet Based Illumination Invariant Preprocessing in Face Recognition. In *Image and Signal Processing, 2008. CISPA ’08. Congress on*, volume 3, pages 421–425, 2008.
- [22] M. Harrah. gedcom4j: A Java library for processing GEDCOM genealogy files. , June 2012. <http://code.google.com/p/gedcom4j/>.
- [23] L. Huang, T. Xia, J. Wan, Y. Zhang, and S. Lin. Personalized portraits ranking. In *Proceedings of the 19th ACM international conference on Multimedia*, MM ’11, pages 1277–1280, New York, NY, USA, 2011. ACM.
- [24] H. Imaoka, A. Hayasaka, Y. Morishita, A. Sato, and T. Hiroaki. Nec’s face recognition technology and its applications. *NEC Technical Journal*, 5(3):28–33, October 2010.
- [25] C. Jingyu, W. Fang, X. Rong, T. Yuandong, and T. Xiaoou. EasyAlbum: an interactive photo annotation system based on face clustering and re-ranking, 2007.
- [26] W.-C. Kao, M.-C. Hsu, and Y.-Y. Yang. Local contrast enhancement and adaptive feature extraction for illumination-invariant face recognition. *Pattern Recognition*, 43(5):1736–1747, 2010.
- [27] Y. J. Lee and K. Grauman. Face discovery with social context. In *Proceedings of the British Machine Vision Conference*, pages 36.1–36.11. BMVA Press, 2011. <http://dx.doi.org/10.5244/C.25.36>.
- [28] Z. Lei, C. Longbin, L. Mingjing, and Z. Hongjiang. Automated annotation of human faces in family albums, 2003.
- [29] D. Little, S. Krishna, J. Black, and S. Panchanathan. A Methodology for Evaluating Robustness of Face Recognition Algorithms with Respect to Variations in Pose Angle and Illumination Angle. In *Acoustics, Speech, and Signal Processing, 2005. Proceedings. (ICASSP ’05). IEEE International Conference on*, volume 2, pages 89–92, 2005.

- [30] S. X. M. Shao and Y. Fu. Genealogical face recognition based on ub kinface database. In *Proc. IEEE CVPR Workshop on Biometrics (BIOM)*, 2011.
- [31] D. Marc, K. Simon, G. Nathan, and S. Risto. From context to content: leveraging context to infer media metadata, 2004.
- [32] M. Naaman, S. Harada, Q. Wang, H. Garcia-Molina, and A. Paepcke. Context data in geo-referenced digital photo collections. In *Proceedings of the 12th annual ACM international conference on Multimedia, MULTIMEDIA '04*, pages 196–203, New York, NY, USA, 2004. ACM.
- [33] M. Naaman, R. B. Yeh, H. Garcia-Molina, and A. Paepcke. Leveraging Context to Resolve Identity in Photo Albums, 2005.
- [34] A. V. Nefian and M. H. Hayes III. Hidden Markov models for face recognition. In *Acoustics, Speech and Signal Processing, 1998. Proceedings of the 1998 IEEE International Conference on*, volume 5, pages 2721–2724 vol.5, 1998.
- [35] W. Ng, T.-M. Zheng, P. Chan, and D. Yeung. Social relationship discovery and face annotation in personal photo collection. In *Machine Learning and Cybernetics (ICMLC), 2011 International Conference on*, volume 2, pages 631 –637, july 2011.
- [36] The Church of Jesus Christ of Latter-day Saints. The GEDCOM Standard Release 5.5, July 2012. http://homepages.rootsweb.ancestry.com/_pmcbride/gedcom/55gctoc.htm.
- [37] A. Pentland, B. Moghaddam, and T. Starner. View-based and modular eigenspaces for face recognition. In *Computer Vision and Pattern Recognition, 1994. Proceedings CVPR '94., 1994 IEEE Computer Society Conference on*, pages 84–91, 1994.
- [38] P. Phillips. MBGC Still Face Challenge version 2 Preliminary Results. In *MBGC Workshop*, 2009.
- [39] H. Robert. Image enhancement by histogram transformation. *Computer Graphics and Image Processing*, 6(2):184–195, 1977.
- [40] J. L. S. Xia, M. Shao and Y. Fu. Understanding kin relationships in a photo. *IEEE Transactions on Multimedia*, pages –, 2012.
- [41] M. S. S. Xia and Y. Fu. Kinship verification through transfer learning. In *Proc. International Joint Conferences on Artificial Intelligence (IJCAI)*, pages 2539–2544, 2011.
- [42] A. Sato, A. Inoue, T. Suzuki, and T. Hosoi. NeoFace: Development of face detection and recognition engine. *NEC Research & Development*, 44(3):302–306, July 2003.
- [43] M. Shao, S. Xia, and Y. Fu. Genealogical face recognition based on ub kinface database. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2011 IEEE Computer Society Conference on*, pages 60 –65, june 2011.
- [44] A. Shashua and T. Riklin-Raviv. The quotient image: class-based re-rendering and recognition with varying illuminations. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(2):129–139, 2001.

- [45] K. Shimizu, N. Nitta, and N. Babaguchi. Learning people co-occurrence relations by using relevance feedback for retrieving group photos. In *Proceedings of the 1st ACM International Conference on Multimedia Retrieval, ICMR '11*, pages 57:1–57:8, New York, NY, USA, 2011. ACM.
- [46] J. Söderberg and E. Kakogianni. Automatic tag generation for photos using contextual information and description logics. In *Content-Based Multimedia Indexing (CBMI), 2010 International Workshop on*, pages 1–7, 2010.
- [47] Z. Stone, T. Zickler, and T. Darrell. Autotagging facebook: Social network context improves photo annotation. In *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW '08. IEEE Computer Society Conference on*, pages 1–8, june 2008.
- [48] M. A. Turk and A. P. Pentland. Face recognition using eigenfaces. In *Computer Vision and Pattern Recognition, 1991. Proceedings CVPR '91., IEEE Computer Society Conference on*, pages 586–591, 1991.
- [49] N.-S. Vu, H. M. Dee, and A. Caplier. Face recognition using the POEM descriptor. *Pattern Recognition*, 45(7):2478–2488, 2012.
- [50] G. Wang, A. Gallagher, J. Luo, and D. Forsyth. Seeing people in social context: recognizing people and social relationships. In *Proceedings of the 11th European conference on Computer vision: Part V, ECCV'10*, pages 169–182, Berlin, Heidelberg, 2010. Springer-Verlag.
- [51] S. Wright. Coefficients of Inbreeding and Relationship. *The American Naturalist*, 56(645):330–338, 1922.
- [52] P. Wu and F. Tang. Improving face clustering using social context. In *Proceedings of the international conference on Multimedia, MM '10*, pages 907–910, New York, NY, USA, 2010. ACM.
- [53] Z. Xiaozheng, G. Yongsheng, and M. K. H. Leung. 3D face modeling using image warping in pose-invariant face recognition. In *Control, Automation, Robotics and Vision, 2002. ICARCV 2002. 7th International Conference on*, volume 1, pages 497–501 vol.1, 2002.
- [54] G. Yan, J. Li, J. Li, Q. Ma, and M. Yu. Illumination Variation in Face Recognition: A Review. In *Intelligent Networks and Intelligent Systems, 2009. ICINIS '09. Second International Conference on*, pages 309–311, 2009.
- [55] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld. Face recognition: A literature survey. *ACM Comput. Surv.*, 35(4):399–458, 2003.