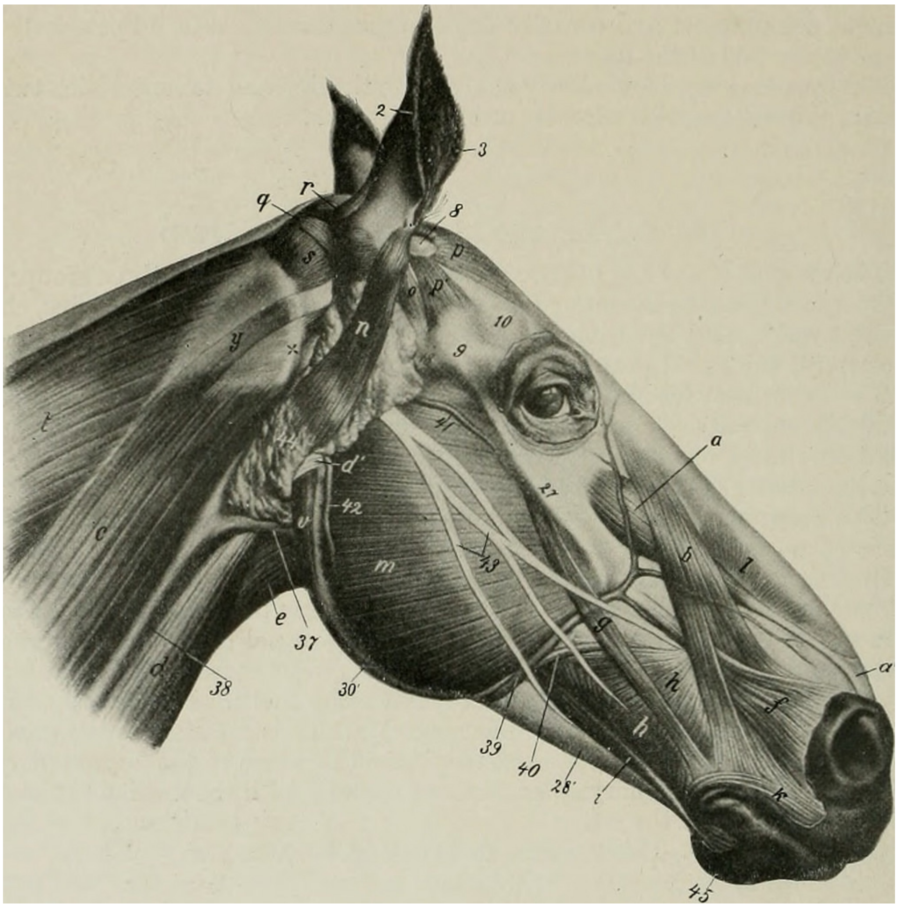
CHALMERS



Automated Detection of Pain in Horses through Facial Expression Analysis

Master's thesis in Complex Adaptive Systems

KASHIF BHATTI

Department of Applied Mechanics
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2016

MASTER'S THESIS IN COMPLEX ADAPTIVE SYSTEMS

Automated Detection of Pain in Horses through Facial Expression Analysis

KASHIF BHATTI

Department of Applied Mechanics

CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2016

Automated Detection of Pain in Horses through Facial Expression Analysis
KASHIF BHATTI

Master's thesis 2016:86
ISSN 1652-8557
Department of Applied Mechanics
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Cover:

Ellenburger-Baum (1914) *Muscles of Head of Horse; Lateral View. The M. cutaneus is Removed* In: Sisson, Septimus, *The anatomy of the domestic animals. p. 256.* Philadelphia and London: W. B. Saunders Company.

a, Levator labii superioris proprius; *b*, levator nasolabialis; *c*, brachiocephalicus; *d*, sterno-cephalicus; *d'*, tendons of d; *e*, omo-hyoideus; *f*, dilatator naris lateralis; *g*, zygomaticus; *h*, buccinator; *i*, depressor labii inferioris; *k*, orbicularis oris; *l*, lateralis nasi, dorsal part; *m*, masseter; *n*, parotido-auricularis; *o*, zygomatico-auricularis; *p*, interscutularis; *p'*, fronto-scutularis, pars temporalis; *q*, cervico-auricularis profundus major; *r*, cervico-auricularis superficialis; *s*, obliquus capitis anterior; *t*, splenius; *v*, occipito-mandibularis; *y*, mastoid tendon of brachiocephalicus; *2*, posterior, *3*, anterior, border of external ear; *8*, scutiform cartilage; *9*, zygomatic arch; *10*, orbital fat; *18*, temporo-mandibular articulation; *27*, facial crest; *30'*, angle of jaw; *37*, external maxillary vein; *38*, jugular vein; *39*, facial vein; *40*, parotid duct; *41*, transverse facial vein; *42*, masseteric vein; *43*, facial nerve; *44*, parotid gland; *45*, chin; *x*, wing of atlas. By an oversight the superior buccal branch of the facial nerve is shown crossing over instead of under the zygomaticus.

Chalmers Reproservice
Göteborg, Sweden 2016

Automated Detection of Pain in Horses through Facial Expression Analysis
Master's thesis in Complex Adaptive Systems
KASHIF BHATTI
Department of Applied Mechanics
Chalmers University of Technology

Abstract

A method for automated pain-assessment in horses through facial-expression analysis is proposed. The method is based on supervised linear classification of a feature stack of Gabor filters and has the desirable quality of not requiring expert knowledge or specialized equipment to make an assessment. The method is evaluated by applying it to images of horses from two clinical trials where the horses were (ethically) subjected to pain. The resulting accuracy of 78% compares favorably to an alternate method of pain assessment based on facial expression cues that requires expertise to administer.

Keywords: facial-expressions, machine-learning, horse, equine, artificial intelligence, pain assessment

Acknowledgements

I thank Krister Wolff for connecting me with the opportunity to work on this project and for taking on the role of examiner.

I am grateful to Pia Andersson (Swedish University of Agricultural Sciences) and Emanuella Dalla Costa (Università degli Studi di Milano) for their excellent clinical investigations on equine pain expression, for allowing me access to their photographic data without which this thesis would not be possible and for their valuable advice and consultation.

Contents

1 Introduction

Pain symptomizes many medical conditions and its presence can have a powerful negative influence on an animal's wellbeing and quality of life. Unless they are prevented from doing so due to factors like disability, adults are able to communicate the presence of pain and the qualities associated with it—such as duration, location and intensity—with clarity and specificity. The challenge of identifying pain in human infants is significantly greater, but even there, humans are able to detect the presence of distress through an innate understanding of vocalizations, facial expressions and other associated indicators of pain.

Detecting pain in animals presents a much larger challenge. Without training and experience, it is not possible to reliably detect the presence of pain in animals. This thesis proposes an approach to detecting pain in large farm animals, applies it to horses and evaluates its performance compared to trained humans.

1.1 Background

The ability to quickly and inexpensively detect pain in large animals has implications for ethics, animal welfare and veterinary medicine. The means of detecting pain in farm animals found in the scientific literature tend to focus on physiological and behavioral changes. Training and experience of the individual carrying out the assessment can influence the results of the evaluation as can his/her knowledge of the species, breed and individual animal. Human empathy and emotions can also effect the results of the evaluation [1]. Methods of pain detection that are automated and do not suffer from the aforementioned drawbacks can be of value.

Facial expressions of pain are commonly used indicators of acute procedural pain in clinical and research environments [2]. Similarities exist between detecting pain in animals and in human infants as both may not be able to communicate feelings of pain and both may display behavioral and psychological changes due to a variety of reasons besides pain. These could include, illness, dejection, general distress etc. Pal, Pritam et al. [3], and others, have been successful in devising systems that can identify pain in infants using images and this motivates an exploration of image-based approaches for animals.

Veterinary literature exists that catalogs and correlates the presence of equine facial expressions with pain. Gleerup et al., [4] have documented behavioral and physiological expressions of pain in six adult horses exposed to to two noxious stimuli; E Dalla Costa et al. [5] have developed a three-step Horse Grimace Scale based on clinical study of forty-six adult horses (of which six formed the control group).

1.2 Goals

The thesis aims to develop and evaluate the performance of an image-based automated system for detecting pain based on observations made by and empirical data collected by Gleerup et al., E Dalla Costa et al. and similar trials.

Recommendations will be made on future work in this direction, and improvements will be suggested to clinical trial procedures and cataloging efforts to facilitate future research efforts.

2 Theory

2.1 Facial Expression Coding

The universality of facial expression in animals has been demonstrated by Darwin in 1872. One of his claims was that there are specific inborn emotions which originated in serviceable associated habits [6]. These are inherited behaviors that are useful responses to certain mental states, such as the furrowing of the eyes to prevent too much light from entering or raising of eye-brows when accessing memory, which is associated with searching ones surroundings where raised eye-brows are useful in increasing one's field of view.

Before attracting broader attention, research interest in human face identification and facial expression detection existed primarily amongst psychologists. Ekman and Friesen conducted seminal research on human facial expression analysis, postulating, in 1971 [7] six prototypic emotional displays consisting of happiness, sadness, fear, disgust, surprise and anger that are universal across human ethnicities and cultures (as cited in [8]).

Ekman and Friesen went on to develop a comprehensive **Facial Action Coding System (FACS)** that encodes and classifies all human facial expressions and associates them with the six prototypic emotional displays [9]. Expressions are broken down into groups of facial muscle actions that produce a feature on the face, such as raised inner eye-brow or tightened lips. These components are called **Action Units (AUs)**. FACS is heavily used in detecting and analysing human facial expressions, both manually and automatically; and in recreating them—such as through computer animations.

Facial expression in non-human animals have been studied to a markedly lesser extent. The pain face of mice [10], horses [11, 5] and primates [12] have been investigated. No universal coding system for any non-human animal has so far been developed.

Dalla Costa et al. have developed a Horse Grimace Scale [5] consisting of 6 AUs for pain assessment. Each AU is scored with three levels of intensity from ”not present”, through ”moderately present” to ”obviously present” (intensity scores 0-3). This may be contrasted against the five scale intensity scoring used in human FACS.

2.2 Automation in Facial Expression Analysis

The availability of inexpensive and easy access to large amounts of computational power has spurred interest in automatic human face detection and to a lesser degree facial expression analysis. Most systems today attempt to classify facial expressions into FACS AUs. So far only marker based systems are able to reliably detect all FACS AUs (citation), although image and video based approaches such as Ref. [13] are able to detect a subset of AUs with high reliability.

Very little research effort has been directed towards automation in detecting animal facial expressions. One such instance is the Rodent Face Finder® developed by Susana G Sotocinal et. al. [10] which automatically detects mice eyes and ears in still-images using boosted cascades of Haar classifiers and estimates the bounds

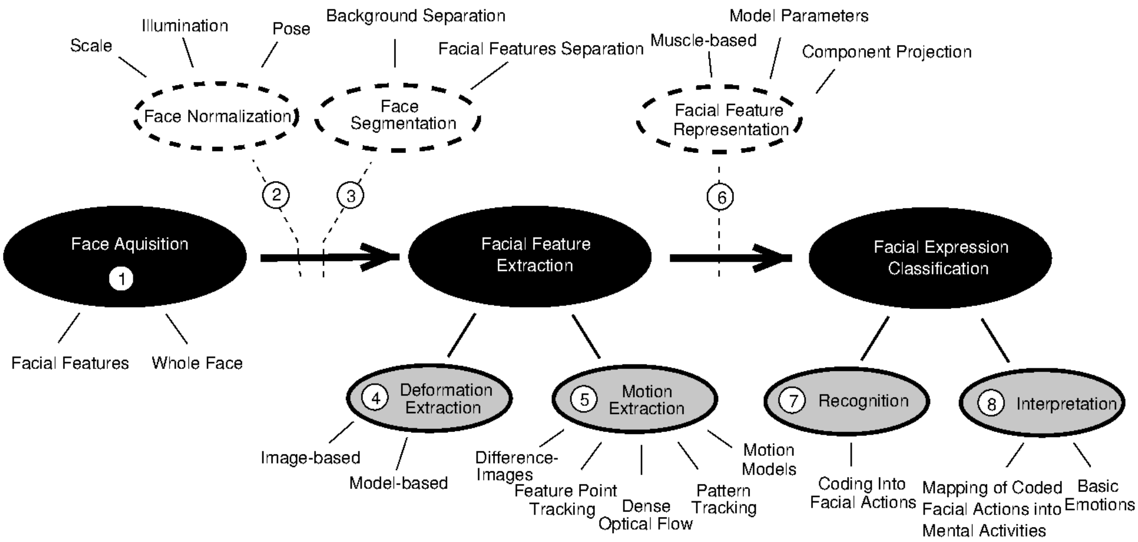


Figure 2.1: *A generic framework for facial feature analysis (reproduced with permission from Ref. [8])*

of their face using this information. This automates only the Face Segmentation stage (see Chapter 2.3) of the process and the remaining stages in the analysis are performed manually.

2.3 An Overview of Automated Facial Expression Analysis

Facial expression analysis is a multi-staged process. Researches work in this area have, over time, developed a variety of tools and techniques for each of these stages. For completeness, and because these concepts serve as a basis for this work, a brief summary of these stages is presented below. Fasel and Luettin [8] have systematized facial expression analysis in the form of a generic framework (see Figure 2.1).

2.3.1 Face Acquisition

The first step in an automated facial expression analysis framework is inevitably the acquisition of a digital representation of a face. The two obvious and commonly used vision based choices are image and video. Other forms are also possible such as displacement data, 3D models, heat maps, marker offsets and so on. This thesis focuses on images of horse faces as input data. This choice is motivated primarily by practical considerations. It has the added advantage of facilitating direct comparison of results with expert human observers, since studies assessing the accuracy of human experts make use of image data.

2.3.2 Normalization

The complexity of detecting facial expression require that as many extraneous variables as possible be eliminated from the analysis. Input data is normalized by making them invariant to, for example, rotation, scale, pose, illumination and/or color.

2.3.3 Segmentation

Before proceeding to feature extraction, some solutions separate the face from the background or isolate *transient* facial features (facial features that result from an active facial expression such as wrinkles caused by

	Holistic methods	Local methods
Deformation extraction		
Image-based	Neural Network	Intensity profiles
	Gabor wavelets	High gradient components
		PCA + Neural Networks
Model-based	Active appearance model	Geometric face model
	Point distribution model	Two view point-based models
	Labeled graphs	
Motion extraction		
Dense optical flow	Dense flow fields	Region-based flows
Motion models	3D motion models	Parametric motion models
	3D deformable models	3D motion models
Feature point tracking		Feature tracking
Difference-images	Holistic diff.-imgs	Region-based difference-images
Marker-based		Highlighted facial features
		Dot markers

Table 2.1: An overview of prominent facial feature extraction methods (reproduced with permission from Ref. [8])

a smile or a blink) and *in-transient* facial features (facial features that are always present on an individual's face, such as shape of nose or wrinkles due to age). This optional stage is particularly common in rule-based approaches.

2.3.4 Feature Extraction

This step involves distilling, through a variety of techniques, a large amount of raw data into a set of values that are appropriate as an input to a machine learning algorithm. Table 2.1 lists some common feature extraction methods found in the literature.

2.3.5 Facial Expression Classification

In this stage, a classifier is employed to reveal AUs, or facial expression cues encoded through other means, associated with the feature set of each face. Neural networks, support vector machine and hidden Markov models are examples of classifiers commonly used during this phase. In addition, before classification is attempted, dimensionality reduction techniques, most commonly principal component analysis (PCA) and Fisher's linear discriminant analysis (LDA), are applied on the feature vector to reduce its size and in some cases improve the classifier's performance.

2.4 Feature Extraction Techniques

2.4.1 Active Appearance Models

Active Appearance Models (AAMs) are generative statistical models of the shape and grey-level appearance of objects and associated fitting algorithms that were first introduced by Cootes et al. [14] and later refined

(a) *In a training set, labeling is typically done manually* (b) *Fitting results for a test image*

Figure 2.2: *Examples of AAM applied to human face. (Adapted from "A Two Step Face Alignment Approach Using Statistical Models" [16] and licensed under CC BY 3.0)*

by Matthews et al. [15] and others. This technique has applications in visually detecting and classifying topologically invariant objects such as human faces and biological organs.

In its original manifestation, an AAM model is generated by fitting *shape* and *appearance* aspects of a face (or other visual entity of interest) to training data using Principal Component Analysis (PCA). Faces that converge to the model are representative of the modeled class. Parameters from the convergent model instance may be used for other purposes such as input to a classifier for facial recognition.

Modeling shape

Consider an $x \times w$ matrix of grey-scale values, A, representing an image of a face. The shape aspect of an AAM for the image is modeled by a mesh of v vertices $\mathbf{s}$.

$$\mathbf{s} = [x_1, y_1, x_2, y_2, \ldots, x_v, y_v]^T \tag{2.1}$$

Typically this mesh is generated by manually marking an identical number of salients points on every image that the model will be trained on or applied to. Each mesh can be expressed as a linear combination of a base shape $\mathbf{s}_0$, the average of all meshes in the training images, and n shape vectors s_i^*.

$$\mathbf{s}_0 = \frac{1}{t}\sum_{i=1}^{t} s_i \tag{2.2}$$

$$\mathbf{s} = \mathbf{s}_0 + \sum_{i=1}^{n} p_i \mathbf{s}_i^* \tag{2.3}$$

PCA of Shape Mesh

Cootes *et al.* apply PCA on the training data to find a pair-wise uncorrelated basis set $\{\mathbf{s}_1^*, \mathbf{s}_2^*, \ldots, \mathbf{s}_n^*\}$ to use in the AAM model.

First the covariance matrix $C \in \mathbb{R}^{2v\times 2v}$ is computed using $\Delta\mathbf{s}_i = \mathbf{s}_i - \mathbf{s}_0$.

$$C_{ij} = E[(\mathbf{s}_i - \mathbf{s}_0)(\mathbf{s}_j - \mathbf{s}_0)] \tag{2.4}$$

$$C = \frac{1}{t}\sum_{i=1}^{t} \Delta\mathbf{s}_i \Delta\mathbf{s}_i^T \tag{2.5}$$

A linearly independent basis for C may now be determined through eigenvalue decomposition by computing Q such that $C = P\Lambda P^{-1}$. Here the i-th column in P is p_i and Λ is a diagonal matrix of eigenvalues λ_i. It is known that for symmetric matrices of real values, the eigenvectors are orthogonal, therefore since $C_{ij} = C_{ji}$, $C = P\Lambda P^T$. The vectors with the largest values in this orthogonal basis constitutes the principal components of the shape mesh.

Modeling Appearance

A similar approach is used to model the appearance of faces i.e. the grey-level textural information embedded in each image.

First all pixels that lie outside of the boundaries of the facial mesh are discarded from A and the resulting matrix vectorized, yielding $\bar{\mathbf{A}}$. This may also be represented as a linear combination of a base shape and m appearance vectors $\bar{\mathbf{A}}_i^*$

$$\bar{\mathbf{A}} = \bar{\mathbf{A}}_0 + \sum_{i=0}^{m} \lambda_i \bar{\mathbf{A}}^* \tag{2.6}$$

Similarly, $\bar{\mathbf{A}}^*$ may be computed through PCA on the training texture data.

Fitting algorithms

A test image may be described by an AAM by locating the closes model instance in its parameter space (p_i and λ_i). The search for this instance is an optimization problem is solved using gradient descent like methods.

2.4.2 2D Gabor Filters

A Gabor filter can be expressed as a complex sinusoidal, referred to as the carrier modulated by a Gaussian shaped function, called the envelope.

$$g(x, y) = \underbrace{s(x, y)}_{\textbf{carrier}} \; \underbrace{w(x, y)}_{\textbf{envelope}} \tag{2.7}$$

The following parametric form is used when implementing the solution.

$$\begin{gathered} g(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \\ x' = x\cos\theta + y\sin\theta \\ y' = -x\sin\theta + y\cos\theta \end{gathered} \tag{2.8}$$

where λ is the wavelength of the sinusoidal factor, θ is the orientation of the normal to the parallel stripes of a Gabor function, γ is the spatial aspect ratio, σ is the standard deviation of the Gaussian envelop, and ψ is the phase offset.

Gabor filters work in the same was as the vision system in humans and animals (improve this text and add a reference) [17]. Unlike some classic approaches such as eigenfaces, Gabor filters can also be made image illumination invariant.

(a) *Original* (b) *Gabor filter*

Figure 2.3: *Gabor convolution applied to a sample image*

3 Method

3.1 Challenges

Human Focus of Existing Research

Research in the area of animal facial expression detection is limited as reviewed in Section 2.2. While inspiration may be drawn from the more active field of human facial expression detection, care must be taken to address the anatomical idiosyncrasies of horses: ears that can point in different direction, fur, patches of colors, lack of a well defined frontal view and so on.

Cataloged Experimental Data

Researchers interested in automation of human facial detection and facial expression analysis have at their disposal various large, high-quality, open-access, labeled data-sets such as the Japanese Female Facial Expression Database (JAFFE) [18], the AR database [19], and the Cohn-Kanade AU-Coded facial expression database [20]. The availability of these databases relieve researchers from the time-consuming, administration-heavy task of collecting and cataloging data. Open-access also eases verification of results and performance comparisons.

As no comparable database for horses is available, a labeled catalog must first be prepared based on input from the two studies discussed in the next section (Section 3.2).

Universal Vocabulary for Equine Facial-Expressions

Pioneering work by Ekman and Friesen (see Section 2.1) has let to a well-developed and widely adapted vocabulary of human AUs. No such comprehensive vocabulary exists for any non-human animal and these AUs cannot be translated to them. Limited AU vocabularies, in the form of grimace scales, for pain have recently been developed for rodents, rabbits and horses (see Section 2.1), however their adaption so far is limited.

3.2 Experimental design of clinical trials and image acquisition

The image data for this study has been collected during two clinical trials on horses. The experimental design of these trials is briefly summarized in Table 3.1, accompanied by details of the facial expression data collected and the conditions under which they were collected.

	Equine pain face *Gleerup et al.*	Horse Grimace Scale *Dalla Costa et al.*
Number of horses	5	40 (+6 control)
Horse gender distribution	Unknown	All stallions (2 mare, 4 gelding in control group)
Pose	Profile	Profile + frontal
Number of pictures	Unknown (6 available)	126 (115 available)
Means of pain induction	Mechanical pressure and capsaicin cream	Castration
Ethical control	Experimental protocol approved by Danish Animal Experiments Inspectorate	Statement of compliance with European directive on protection of animals used for experimental purposes (86/609/EEC). Registered with Brandenburg State Veterinary Authority

Table 3.1: Overview of clinical trials

3.2.1 An equine pain face

Gleerup et al. [11] have conducted a clinical trial on five healthy adult horses (a sixth horse was injured on pasture and excluded). Pain was induced externally using two different noxious stimuli: mechanical pressure through a tourniquet on the antebrachium, and neurogenic inflammation through topical application of capsaicin cream.

In total, six trials were conducted on each horse. Pain was applied through mechanical means for two trials, pain was applied through topical application of capsaicin cream for two trials, and for control, no noxious stimulus was applied for two trials. One trial in each pair of trials was conducted with a human observer present (and visible to the horse) for the duration of the trial and one without.

Pictures and videos of the horses, with particular focus of their faces, were recorded during the trials.

3.2.2 Development of the Horse Grimace Scale (HGS)

Emanuela Dalla Costa et al. [5] have developed 6 facial expression AUs (see Figure 3.1) that are associated with pain in horses and conducted trials to assess and validate their effectiveness. The trial consists of forty six horses divided into groups A, B and C consisting of 19, 21 and 6 horses respectively. Groups A and B underwent surgical castration, while Group C was subjected to non-invasive procedures associated with little or no pain induction (x-ray, hoof corrections and teeth rasping) and served as a control group. While group A and B were obviously entirely male, group C consisted of two meres (adult female horses) and four geldings (castrated males) for practical reasons. The average age of the control group was 5.5 years, while the average age of the test groups was 2.3 years. The researchers have attempted to minimize the influence of behavioral difference between genders and ages by providing an acclimatization period of two days to all horses before the trial.

Group A received an analgesic 5 minutes before being anesthetized, group B received an additional dose of analgesic 6 hours post-op, and group C received only the pre-op analgesic for control. All three groups were anesthetized under the standard protocol for general anesthesia in horses. Video recording were made before the procedure and 8 hours post-op. Favorable stills were extracted from these videos and cropped retaining only the head and these were scored on the HGS by a blind group of evaluators who were experts either in horses or facial expression scoring.

Figure 3.1: *AUs comprising the Horse Grimace Scale as developed by Dalla Costa et al. (Reproduced from Ref. [5] and licensed under CC BY 4.0)*

Group A	19 stallions	Castration	General anesthesia Analgesic 5 minutes prior to procedure
Group A	21 stallions	Castration	General anesthesia Analgesic 5 minutes prior to procedure Analgesic 6 hours after procedure
Group A	2 mares, 4 geldings	Castration	General anesthesia Analgesic 5 minutes prior to procedure

Table 3.2: Horses involved in the Horse Grimace Scale study were divided into three groups as shown.

3.3 System design

This thesis focuses on deformation based facial feature extraction of available image data and its interpretation. This corresponds to stages (4) and (8) in Fasel and Luettin's generic framework (Figure 2.1).

3.3.1 Choice of pose

A good choice of pose for automated facial expression analysis will facilitate capturing of as much facial expression information as possible. In humans this choice is obvious and indeed a common characteristic of all (non-pose invariant) human facial recognition studies is that they rely on the frontal face view.

Due to the anatomy of a horse's head, a frontal projection has a smaller area than a side projection. The narrow, elongated profile obscures details around the mouth, cheeks and eyes. This frontal view is dominated by a large forehead that runs down to the muzzle and, unlike humans, offer no expressiveness. These factors make a frontal view less than ideal for capturing facial expression details.

A profile view may be a superior choice as it captures more expressive facial features in plane. Not all facial expressions are symmetric, however, and the asymmetry of the facial expressions will contain valuable information which is lost if only one profile view is used. This may be addressed by taking multiple profile views or combining a profile views with a frontal view.

The same pictures were used for the purpose of evaluating the automated system proposed by this thesis and used by human evaluators to assess pain in both clinical trials. In both trials this is a slightly rotated profile view. This captures details in and around the jaw-line, cheek, nostrils and eyes, and even captures any asymmetry in ear positions that may be missed or obscured in a purely profile view. This choice has the added advantage of leveling the playing field so that meaningful comparisons can be made between the performance of the proposed automated system and that of human experts. Of course information on asymmetry of facial features besides those already mentioned is not captured. Figure 3.2 illustrates this pose.

3.3.2 Preprocessing

Before feature generation and training can be carried out, several manual and automated steps are performed to prepare the images for use. The goal is to prepare data for use by feature extraction algorithm and circumvent the need for implementing an automated system for face detection, location, scaling and orientation, which is not the focus of this study.

Screening

Training images are screened and those images that are determined to be too poor in quality are removed from the training-set. Images with obvious image-quality defects such as poor focus or over-exposed images are excluded at this stage. Other factors that may affect the classification algorithm or feature computation may also result in exclusion, such as obstruction of important facial features, incorrect head pose etc.

Landmark labeling

Three salient facial key-points or landmarks are labeled on each image. The location of these points were chosen carefully to allow reliable labeling even on low-quality images. They are located at the base of nostril, the center of eye, and at the apex of the crease formed by the chewing muscles (see Figure 3.2).

Figure 3.2: *Profile view of horse with anchor-points market*

Normalization

Images are mirrored if necessary and a three-point affine transformation, based on the three manually labeled landmarks, is applied to canonicalize each image.

3.3.3 System design

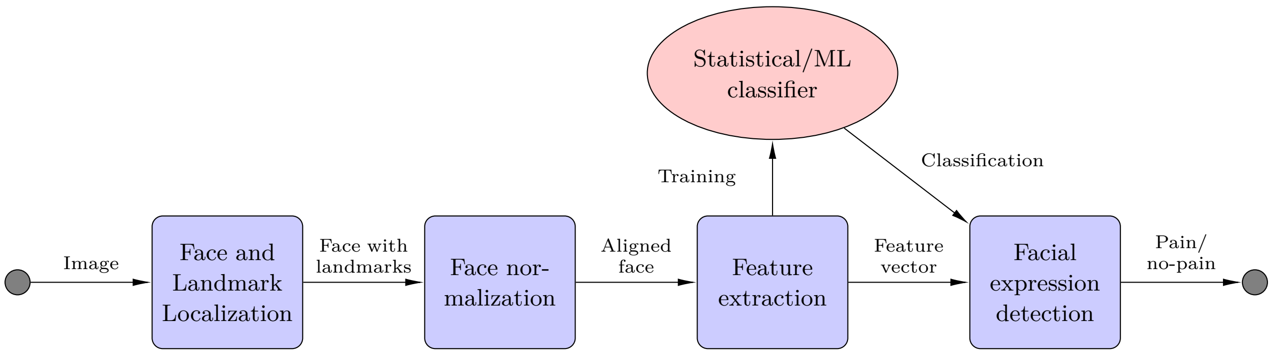


Figure 3.3: *Flow chart for facial expression detection. Modeled on generalized work-flow presented in Ref. [21].*

Based on factors discussed above and in the preceding theory section, a proposed high-level design for a system to detect Equine facial-expressions is presented in Figure 3.3.

3.4 Feature extraction

There are a number of options available for featurizing facial expression information from static images. Table 2.1 lists some common approaches and Section 2.4 describes two of them. For reasons that follow, Gabor Filters were selected as the method of choice for this study.

Since it is known that human observers are able to detect facial expressions of pain in horse faces, a feature that approximates human vision systems is a *safe* choice. Gabor filters are also more robust against occlusion by hair or harness (AAMs are known to be susceptible to performance issues when parts of human faces are occluded with, for example, thick-rimmed glasses [22]). They are known to perform well with low resolution images, and imprecisely located key-points [23]. These advantages are important since most images used in this study are low-resolution stills extracted from video and are hand-labeled. A deciding factor in eliminating AAMs from consideration is that the construction of a face-model requires a large number of training instances and no model for the horse-face is available (precomputed human face models are widely available).

A distinguishing feature of Gabor Filters and similar techniques is that they allow for direct interpretation of results into emotions. This is viewed as an advantage due to the absence of a broadly accepted and validated[1] coding-system for horse-face AUs associated with pain. It also avoid the problem of facial AUs mapping to multiple emotional states.

The feature vector is based on a stack of Gabor convolutions of a local patch of pixel around a set of predetermined key-points. Key points were generated by a uniform grid over the entire image, including background and occluded regions. Each patch is convoluted with 8 orientations $\pi/8, \pi/4, 3\pi/8 \ldots 2\pi$ and 5 spatial frequencies in the range 0.4 and 2.5 for a total of 40 filters per image (Figure 3.4). This approach is adapted from Ou [25] and Rose [23].

θ in (2.8) corresponds to the orientation and the spatial frequency, b can be related to the wavelength λ and the standard deviation σ.

$$\frac{\sigma}{\lambda} = \sqrt{\frac{\log 2}{2}} \frac{2^b + 1}{\pi(2^b - 1)} \tag{3.1}$$

The real part of Equation 2.8 (2.8)

$$g^R(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \exp\left(i\left(2\pi\frac{x'}{\lambda} + \psi\right)\right) \tag{3.2}$$

can then be convolved with the image patch $I_n(x, y)$ to yield the filter.

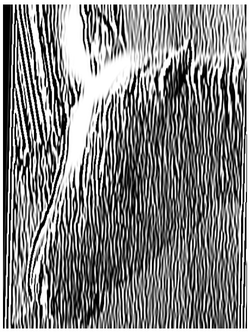
$$O_{1\ldots40}(x, y) = I_n(x, y) * g^R_{1\ldots40}(x, y) \tag{3.3}$$

3.5 Classification

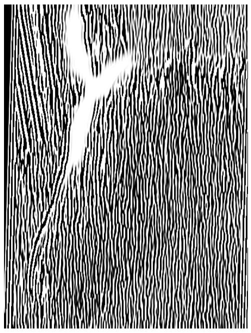
A support vector machine (SVM) is used for classification, due to its robustness when dealing with data of high dimensionality and its performance. The feature vector is classified directly to pain state, without using an intermediate classification to AUs as is common with human facial expression analysis. This is motivated by a lack of well developed FACS for non-human animals and uncertainty around their mapping to emotional classes.

[1] Performing such validation is in itself a challenging task, since emotional states are not uniquely defined by AUs, or through any other means, and their (including Ekman's) descriptions in terms of the aforementioned are ambiguous [24].

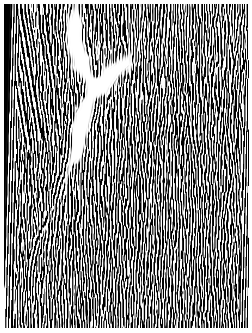
(a) $\theta = 0$, $f = 0.16$



(b) $\theta = 0$, $f = 0.18$



(c) $\theta = 0$, $f = 0.20$

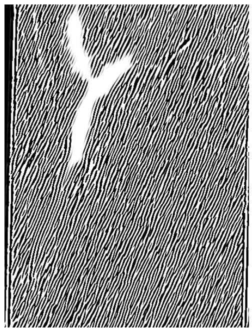


(d) $\theta = 0$, $f = 0.22$

(e) $\theta = \pi/8$, $f = 0.16$

(f) $\theta = \pi/8$, $f = 0.18$

(g) $\theta = \pi/8$, $f = 0.20$



(h) $\theta = \pi/8$, $f = 0.22$

(i) $\theta = \pi/4$, $f = 0.16$

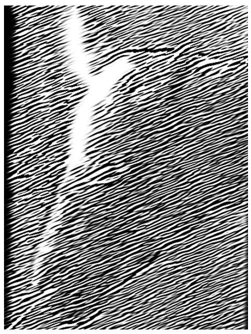
(j) $\theta = \pi/4$, $f = 0.18$

(k) $\theta = \pi/4$, $f = 0.20$

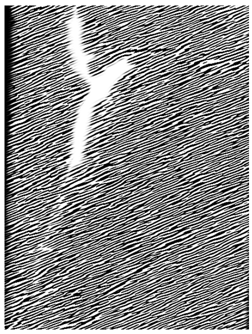
(l) $\theta = \pi/4$, $f = 0.22$

(m) $\theta = 3\pi/8$, $f = 0.16$

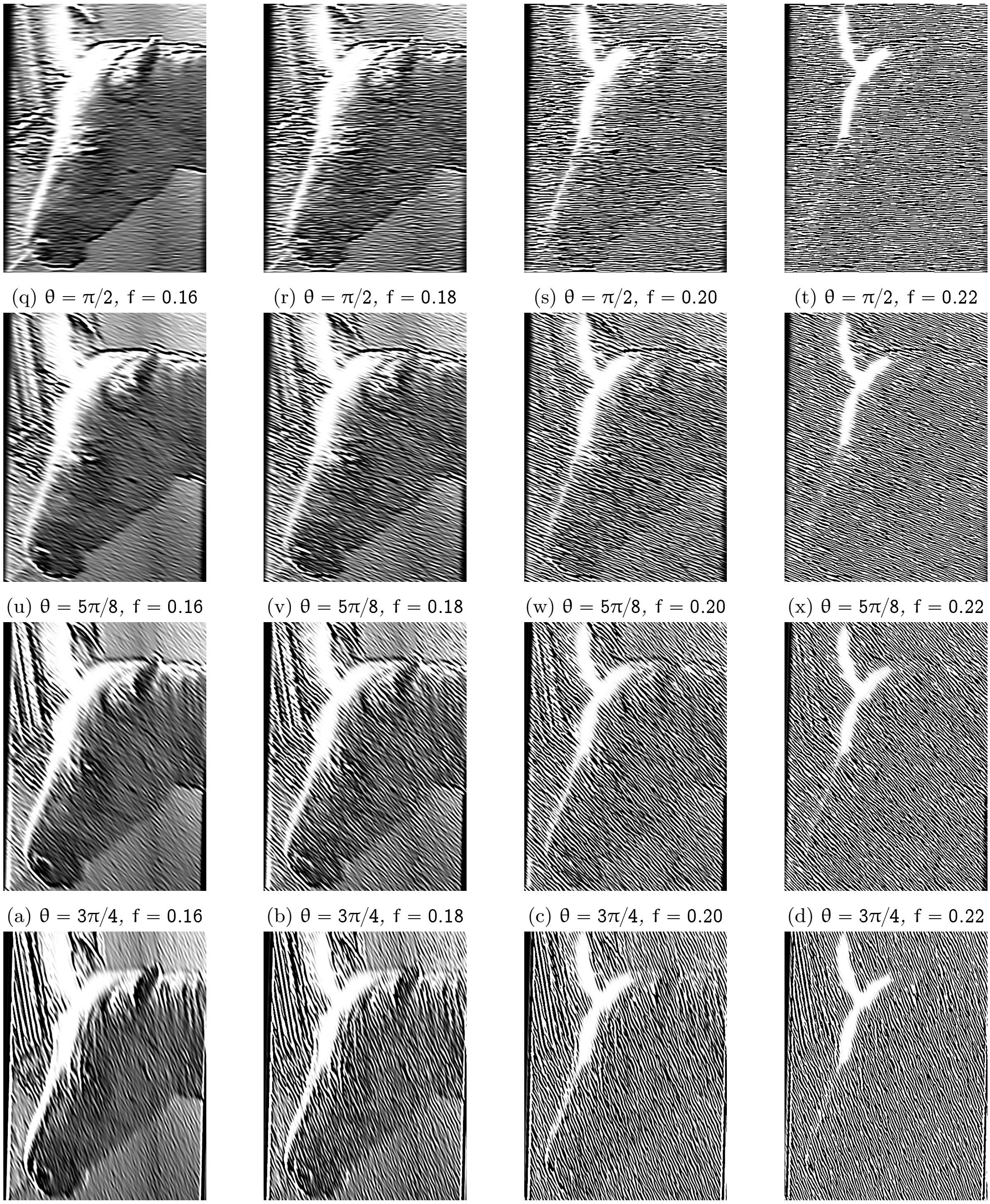
(n) $\theta = 3\pi/8$, $f = 0.18$



(o) $\theta = 3\pi/8$, $f = 0.20$



(p) $\theta = 3\pi/8$, $f = 0.22$



(q) $\theta = \pi/2$, $f = 0.16$ (r) $\theta = \pi/2$, $f = 0.18$ (s) $\theta = \pi/2$, $f = 0.20$ (t) $\theta = \pi/2$, $f = 0.22$

(u) $\theta = 5\pi/8$, $f = 0.16$ (v) $\theta = 5\pi/8$, $f = 0.18$ (w) $\theta = 5\pi/8$, $f = 0.20$ (x) $\theta = 5\pi/8$, $f = 0.22$

(a) $\theta = 3\pi/4$, $f = 0.16$ (b) $\theta = 3\pi/4$, $f = 0.18$ (c) $\theta = 3\pi/4$, $f = 0.20$ (d) $\theta = 3\pi/4$, $f = 0.22$

(e) $\theta = 7\pi/8$, $f = 0.16$ (f) $\theta = 7\pi/8$, $f = 0.18$ (g) $\theta = 7\pi/8$, $f = 0.20$ (h) $\theta = 7\pi/8$, $f = 0.22$

Figure 3.4: *Gabor stack for one of the horse faces generated with* $\sigma = 31$, $\theta = \pi/2$, $\lambda = 0.2$, $\gamma = 0.5$, $\psi = \pi/2$, $\lambda = 1/f$ *for* $\theta = 0 \dots \pi$. *The black strips at the top left and bottom right corners are artificats of the affine-transform.*

4 Results and discussion

The procedure has been applied on a combined set of 64 images from both clinical trials. Using 95% of the images as a test-set, an average prediction accuracy of 78% has been achieved with a small bias towards false positive results (12.5% versus 9.5% for false negatives).

The effect of test-set size on prediction accuracy has been investigated and the results summarize in Figure 4.1. Prediction accuracy is expected to increases with increasing test-set size until it asymptotically approaches a peak value. Since the increase in accuracy is approximately linear for the entire range of test-set sizes tested, we conclude that the classifier has not been adequately trained. An increase in performance is therefore expected with a larger training-set size.

4.1 Comparison to human based trials

Dalla Costa et al. have carried out assessment of the prediction accuracy of the HGS (see Section 3.2.2). Five individuals with expertise in either horses or facial-expression assess and assign intensity scores to AUs described as *stiffly backwards ears*, *orbital tightening*, *tension above the ear area*, *prominent strained chewing muscles*, *mouth strained and pronounced chin*, and *strained nostrils and flattening of the profile*. Figure 3.1 presents examples of these AUs. The values are then used by the experts to make an overall pain assessment. The study reports a prediction accuracy ranging from 67.5 to 77.8%, with an average of 73.3%. 17.0% of incorrect evaluations were false positives and 9.8% false negatives. We observe that raw score (sum of intensity values for all AUs) have poor discrimination power for values between 2 and 5 which corresponds to approximately 33% of the spectrum of values (see Figure 4.3).

While the results of the proposed automated approach compare well with those reported in the study, it is worth noting that there are systematic differences in both performance analyses. Dalla Costa's evaluation includes both front and lateral pictures of horses, while those used in the automated evaluation employed only lateral views; and their pictures were not screened for image-quality, orientation or obstruction, while those used in the automated evaluation were (see Section 3.3.2). The effect of screening is not expected to have a large impact on the performance of human evaluators since human vision is more tolerant to obstruction and out of plane rotation compared a computer algorithm.

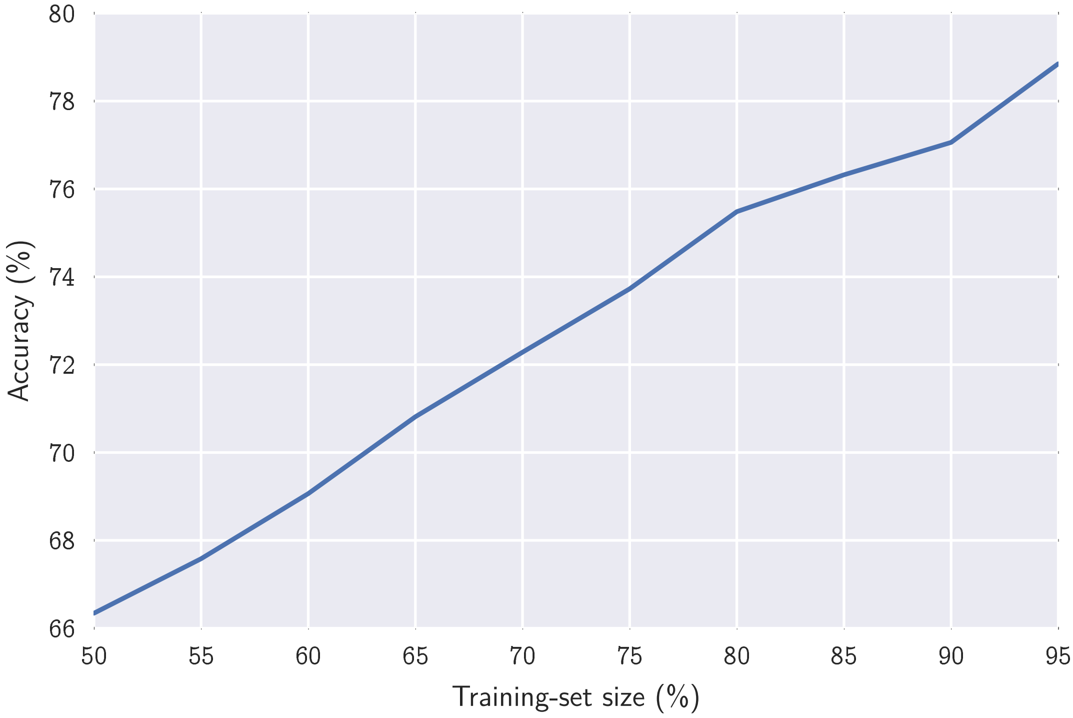


Figure 4.1: *Prediction accuracy plotted against test-set size as a percentage of the maximum possible test-set size (of 64 images) where* $\sigma = 3$*,* $\gamma = 0.5$*,* $\psi = 3\pi/4$*,* $f_{mean} = 0.2$*,* $\Delta f = 0.02$*. Key-points are spaced at 50 pixels in a grid formation. All plotted values are the average of 1000 runs and the images are partitioned into training- and test-sets randomly for each run.*

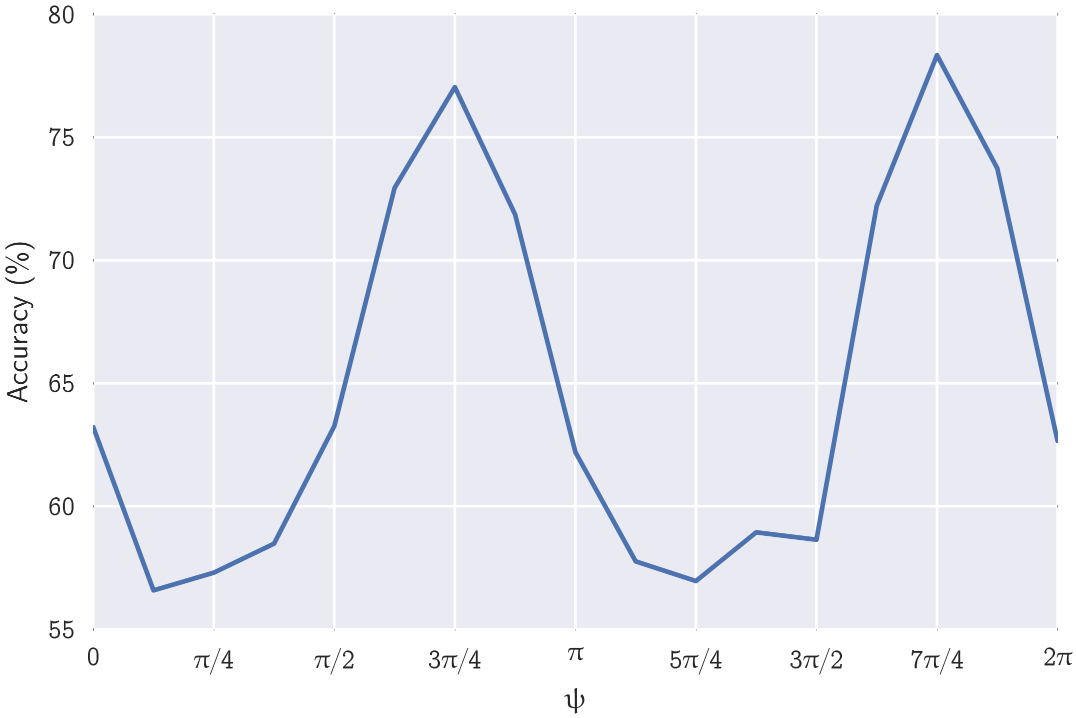


Figure 4.2: *For a grid size of 50 pixels an approximate sinusoidal relation ship between accuracy and* ψ *is observed. Here* $\sigma = 3$*,* $\gamma = 0.5$*,* $f_{mean} = 0.2$*,* $\Delta f = 0.02$*, and 90% of images are used as a training data.*

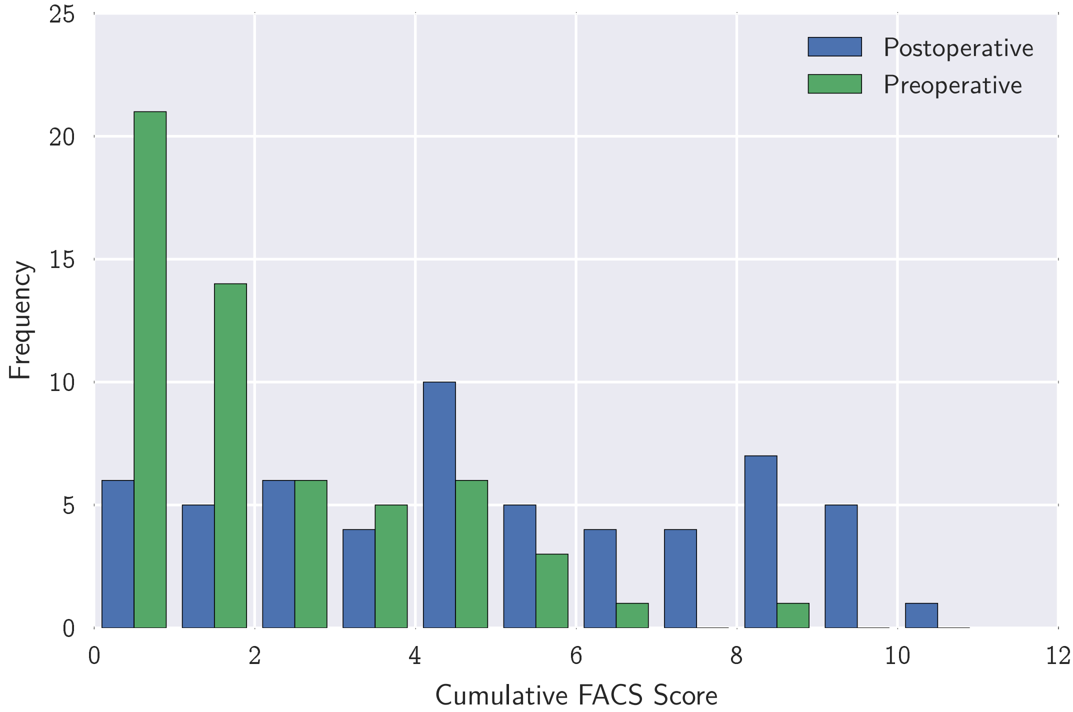


Figure 4.3: *Cumulative AU intensity scores from expert HGS evaluation. Observe that the cumulative score does not discriminate well for values between 2 and 5.*

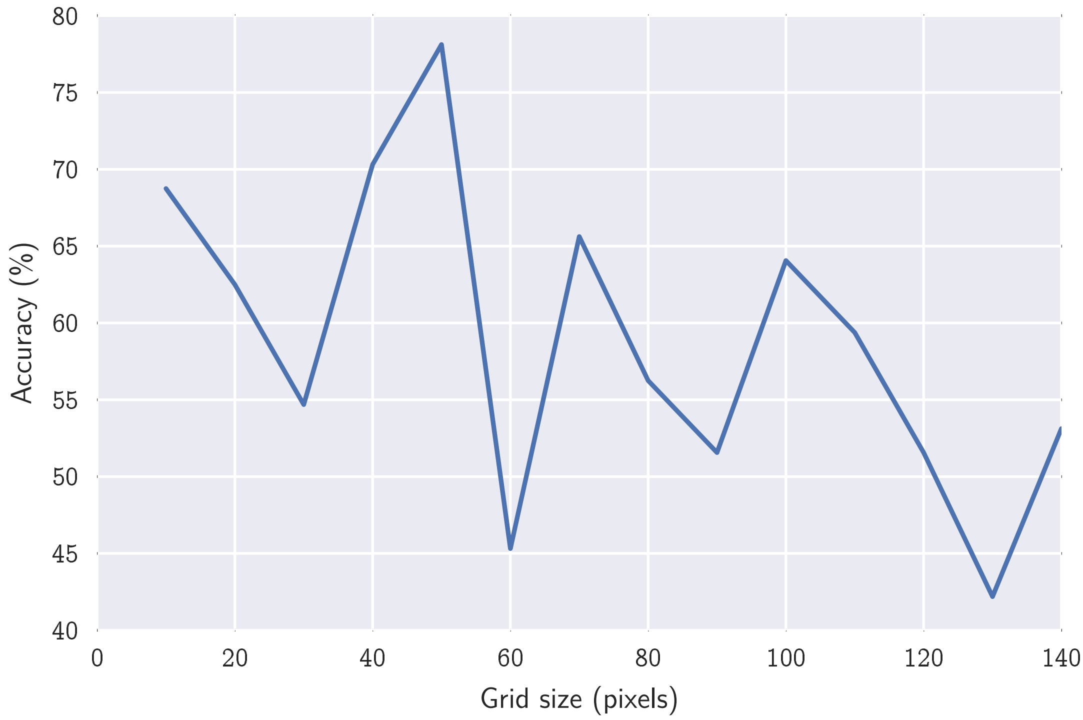


Figure 4.4: *Investigating the effects of grid-size on prediction accuracy. Here* $\sigma = 3$, $\gamma = 0.5$, $f_{mean} = 0.2$, $\Delta f = 0.02$, $\psi = 3/4\pi$, *and all but one available image is used as training-set. The average accuracy of all 64 possible partitioning combinations is plotted for each grid size. The low sample size results in volatile results but a general trend of lower accuracy with increasing grid size for grid size values higher than 100 pixels is noticeable.*

5 Conclusions and recommendations

Based on the preceding findings, detecting pain through automated facial expression analysis in horses is viable. It has potential as an addition to the toolbox of non-invasive equine pain assessment and can compliment expert physiological, behavioral and facial-expression based assessments. The findings of this limited study suggest that it may even offer a superior alternative to an expert evaluated grimace scale, in accuracy, speed and convenience.

The method proposed has the advantage of not requiring specialized equipment (besides cameras which are ubiquitous in this age) or training. Once a classifier has been trained, pain evaluation can be made in a matter of minutes.

Since the proposed automation approach is agnostic to the emotion under assessment, it may have applications for other emotional states. The challenge will be in constructing experiments where images can be captured with the with the emotion under test can be reliably captured. This is easy to achieve with humans, but require careful design and domain specific knowledge.

5.1 Proposals for future work

5.1.1 Animal facial expression database

A key challenge in this thesis has been the lack of a high-quality, expertly labeled image database. Researchers working with human face detection and recognition and human facial expression detection have access to several databases for research purposes (see Section 3.1). There are strong arguments for creating such a database for horses (and indeed other mammals).

Compared to collecting humans facial expression data, compiling labeled data on equine facial recognition is a vastly more complex endeavor, requiring expertise in veterinary sciences, animal behaviorally, clinical research, anesthesiology as well as data and computer sciences.

Such database can allow researchers in the ”analysis” aspect of facial expression analysis—such as codification of facial expressions (into AUs or similar), performing comparative studies between animals, modeling exercises or investigations into automation—to avoid diverting time and resources into this necessary preparatory step. Researchers and practitioners may easily compare results of various techniques by using these databases as benchmarks (this is common with human face and facial expression databases). They also ease verification of results which may be challenging to do otherwise.

For these reasons, we recommend that an image database modeled on the aforementioned databases be developed for animal facial expressions to spur and facilitate research in this area.

5.1.2 Comparative study of automated facial expression approaches

This approach for automated facial expression analysis has been developed around the constraints of available data. It is by no means is the only possible approach. A study evaluating the suitability for automated equine facial-expression analysis of other approaches, such as those mentioned in Table 2.1 and their comparative performance may be a natural next step. The database of animal facial expressions proposed in the preceding section will certainly help.

References

[1] V. Molony and J. E. Kent. Assessment of Acute Pain in Farm Animals Using Behavioral and Physiological Measurements^ 1^,^ 2. *JOURNAL OF ANIMAL SCIENCE-MENASHA THEN ALBANY THEN CHAMPAIGN ILLINOIS-* **75** (1997), 266–272.

[2] A. Desrosiers, D. Harrison, and A. Letham. Use of facial expressions for pain assessment in infants during acute painful procedures. **17**.1 (2015).

[3] N. Streets. EMOTION DETECTION FROM INFANT FACIAL EXPRESSIONS AND CRIES Pritam Pal , Ananth N . Iyer and Robert E . Yantorno , (2006), 721–724.

[4] K. B. Gleerup et al. An equine pain face. *Veterinary anaesthesia and analgesia* **42**.1 (Jan. 2015), 103–14. ISSN: 1467-2995. DOI: 10.1111/vaa.12212. URL: http://www.ncbi.nlm.nih.gov/pubmed/25082060%20http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC4312484.

[5] E. Dalla Costa et al. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. *PloS one* **9**.3 (2014), e92281. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0092281.

[6] C. Darwin, P. Ekman, and P. Prodger. *The expression of the emotions in man and animals*. Oxford University Press, USA, 1998.

[7] P. Ekman and W. V. Friesen. Constants across cultures in the face and emotion. *Journal of personality and social psychology* **17**.2 (1971), 124.

[8] B. Fasel and J. Luettin. Automatic facial expression analysis: a survey. *Pattern Recognition* **36**.1 (Apr. 2003), 259–275. ISSN: 0031-3203. DOI: 10.1016/S0031-3203(02)00052-3.

[9] P. Ekman and W. V. Friesen. Facial action coding system (1977).

[10] S. G. Sotocinal et al. The Rat Grimace Scale: A partially automated method for quantifying pain in the laboratory rat via facial expressions. *Molecular Pain* **7**.1 (2011), 55. ISSN: 1744-8069. DOI: 10.1186/1744-8069-7-55. URL: http://www.molecularpain.com/content/7/1/55.

[11] K. B. Gleerup and C. Lindegaard. Recognition and quantification of pain in horses: A tutorial review. *Equine Veterinary Education* (2015), n/a–n/a. ISSN: 09577734. DOI: 10.1111/eve.12383. URL: http://doi.wiley.com/10.1111/eve.12383.

[12] S. Chevalier-Skolnikoff. Facial expression of emotion in nonhuman primates. *Darwin and facial expression: A century of research in review* (1973), 11–89.

[13] J. Hamm et al. Automated Facial Action Coding System for dynamic analysis of facial expressions in neuropsychiatric disorders. *Journal of Neuroscience Methods* **200**.2 (2011), 237–256. ISSN: 01650270. DOI: 10.1016/j.jneumeth.2011.06.023. URL: http://linkinghub.elsevier.com/retrieve/pii/S016502701100358X.

[14] T. F. Cootes, G. J. Edwards, and C. J. Taylor. Active appearance models. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **23**.6 (2001), 681–685. ISSN: 01628828. DOI: 10.1007/BFb0054760. URL: http://ieeexplore.ieee.org/xpls/abs%7B%5C_%7Dall.jsp?arnumber=927467$%5Cbackslash$nhttp://link.springer.com/chapter/10.1007/BFb0054766.

[15] I. Matthews et al. Active appearance models revisited. *International Journal of Computer Vision* **60**.2 (2004), 135–164. ISSN: 0920-5691. DOI: 10.1023/B:VISI.0000029666.37597.d3. URL: http://www.springerlink.com/index/K1N57K0228074X08.pdf.

[16] Y. Cui, Z. Jin, and W. Yang. A two step face alignment approach using statistical models. *International Journal of Advanced Robotic Systems* **9** (2012), 1. ISSN: 17298806. DOI: 10.5772/52207. URL: http:

//www.intechopen.com/journals/international%7B%5C_%7Djournal%7B%5C_%7Dof%7B%5C_%7Dadvanced%7B%5C_%7Drobotic%7B%5C_%7Dsystems/a-two-step-face-alignment-approach-using-statistical-models.

[17] J. Saragih and R. Göcke. Learning AAM fitting through simulation. *Pattern Recognition* **42**.11 (2009), 2628–2636. ISSN: 00313203. DOI: 10.1016/j.patcog.2009.04.014.

[18] M. J. Lyons et al. The Japanese female facial expression (JAFFE) database (1998).

[19] A. Martinez and R. Benavente. The AR face database, 1998. *Computer Vision Center, Technical Report* **3** (2007).

[20] P. Lucey et al. "The Extended Cohn-Kanade Dataset (CK+): A complete dataset for action unit and emotion-specified expression". *Computer Vision and Pattern Recognition Workshops (CVPRW), 2010 IEEE Computer Society Conference on*. IEEE. 2010, pp. 94–101.

[21] A. K. Jain and S. Z. Li. *Handbook of face recognition*. Vol. 1. Springer, 2005. Chap. 1.

[22] Y.-K. W. Y.-K. Wang et al. Improvement of Face Recognition by Eyeglass Removal. *Intelligent Information Hiding and Multimedia Signal Processing (IIH-MSP), 2010 Sixth International Conference on* (2010). DOI: 10.1109/IIHMSP.2010.64.

[23] N. Rose. "Facial Expression Classification using Gabor and Log-Gabor Filters". *7th International Conference on Automatic Face and Gesture Recognition, 2006. FGR 2006*. IEEE, 2006, pp. 346–350. ISBN: 0-7695-2503-2. DOI: 10.1109/FGR.2006.49. URL: http://ieeexplore.ieee.org/ielx5/10739/33863/01613044.pdf?tp=%7B%5C&%7Darnumber=1613044%7B%5C&%7Disnumber=33863%20http://ieeexplore.ieee.org/xpls/abs%7B%5C_%7Dall.jsp?arnumber=1613044%20http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1613044.

[24] M. Pantic and L. J. M. Rothkrantz. Automatic analysis of facial expressions: the state of the art. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**.12 (2000), 1424–1445. ISSN: 0162-8828. DOI: 10.1109/34.895976. URL: %7B%3C%7DGo%20to%20ISI%7B%3E%7D://WOS:000165901900006$%5Cbackslash$nhttp://ieeexplore.ieee.org/ielx5/34/19391/00895976.pdf?tp=%7B%5C&%7Darnumber=895976%7B%5C&%7Disnumber=19391%20http://ieeexplore.ieee.org/ielx5/34/19391/00895976.pdf?tp=%7B%5C&%7Darnumber=895976%7B%5C&%7Disnumber=19391%20http://ieeexplore.ieee.org/xpl/articleDe.

[25] J. O. J. Ou et al. "Automatic Facial Expression Recognition Using Gabor Filter and Expression Analysis". *Second International Conference on Computer Modeling and Simulation, 2010. ICCMS '10*. Vol. 2. 2010, pp. 215–218. ISBN: 978-1-4244-5642-0. DOI: 10.1109/ICCMS.2010.45. URL: http://ieeexplore.ieee.org/ielx5/5420948/5421056/05421091.pdf?tp=%7B%5C&%7Darnumber=5421091%7B%5C&%7Disnumber=5421056%20http://ieeexplore.ieee.org/Xplore/cookiedetectresponse.jsp.