Project Background and Description

R. S. Schestowitz
Imaging Science and Biomedical Engineering
Stopford Building
University of Manchester
United Kingdom

11th November 2003

Abstract

Landmark selection in appearance models proves to be a complex problem that is presently solved in ways which are by no obvious criterion optimal. Given a set of training data, diffeomorphic warps are used to pose a simpler correspondence problem, thereby generating more compact models of appearance. In recent work, statistical models of intensity and shape are combined in a rather artificial way that ignores the inherent correlation between the two. Furthermore, construction of appearance models has not yet been made a fully automated process. The need for manual annotation seems redundant given some intriguing developments in model-based image analysis and exploitations of these developments bears potential with respect to optimisation and complete automation. This paper attempts to describe the existing problems in some depth and outline some of the currently conceived solutions.

1 Introduction

Shape and appearance models are often used to represent data (visually) whose properties are roughly known in advance. These statistical models are defined to be flexible enough to generalise to different types of legal data examples, yet to preserve some invariants and constraints so that no illegal examples are judged to be acceptable. The way in which such models work is that they deform to fit a feature in a given image and the allowable variance is dictated by some values that are incorporated into the model. These values are derived from what is called a *training* set which defines legal sets of values. This process is the analogue of "teaching" a system how to make sensible decisions. Subsequently, some performance evaluation of the trained system is required to infer its accuracy, for example an unbiased error rate. Much more detail can be found in [1].

The use of such models has been quite successful, but accuracy and speed are still two hot topics. The search for good models continues as more demanding applications of higher resolution and higher dimensionality become available. Another problem worth solving is the automatic annotation of images. Not only can it save valuable time of experts in a field, but it can also help in the acquisition of a large number of reliable, precise, unbiased and inexpensive annotated images. With more data handled without human intervention, more input is available to train classification systems such as shape and appearance models.

Section 2 explains some of the key concepts that later on clarify how active appearance models work. It also introduces some concepts that can aid in solving the existing problems and deficiencies of active appearance models as described in Section 3 onwards. More details are available in the referenced material and only short explanatory notes are provided to keep this document sufficiently broad in scope. The last section summarises the issues and concludes on the measures to be taken to tackle them. No substantial developments or ideas are proposed in this paper as it strives to just provide a general overview and a roundup of the state of existing ASM/AAM "technology".

2 Background

Automatic landmark generation, or more broadly landmark selection, has been an issue of great exploration in the past few years. As one of the ultimate goals of image analysis tasks is complete automation and a precise deterministic approach to selection, older techniques such as manual annotation of an image by experts is a task that ought to be emulated in a reliable way by machine intelligence. Brute-force has been used to enable complex learning tasks for quite some time and investigation points of interest *en masse* in an image, e.g. lines of high curvature, is certain to lead to automatic annotation of some quality. The level of accuracy of such process, however, which is strongly dependent on the algorithms used, still appears to be a major hindrance. Methods of landmark selection and full automation have been described in [2, 3] and more recently a good solution have been discovered by Davies *et al.* [4, 5, 6] for landmark selection in statistical shape models.

Problems associated with dimensionality have been pointed out in the literature above. It is vital to ensure that methods work regardless of the number of dimensions dealt with as one of the strengths of manual annotation only becomes apparent when analysing 3-D data. This is primarily due to the impossibility of annotating a large number of slices manually, as in the case of medical imaging.

The rest of Section 2 attempts to objectively explain some of the more fundamental concepts that build up towards the later developments and proposal of new methods.

2.1 Active Appearance Models

Deformable appearance models of shape and intensity are described in a reasonable level of detail by Cootes *et al.* in [7] and in [8] although they were first introduced by Edwards *et al.* in [9]. The basic idea was that a measure of similarity between image intensities guided a progressive search. Minimisation of the difference between a couple of images using the sum of a pixel-wise comparison brought the statistical appearance model and its target to convergence[1]. During this process, parameters were repetitively being re-evaluated so that they better described the target object – that is the object in the image that resembled the model (see Figure 1). Likewise, free-hand manipulation of these parameters allowed synthesis of new realistic images[2], meaning that a set of assignments for a collection of parameters b_1,$b_2, ...,$ b_n would describe a legal (earthly from an anatomical perspective) type of variation for the model. The following figure shows an appearance model hovering over some target. The model is highlighted in red and labelled 'M'.

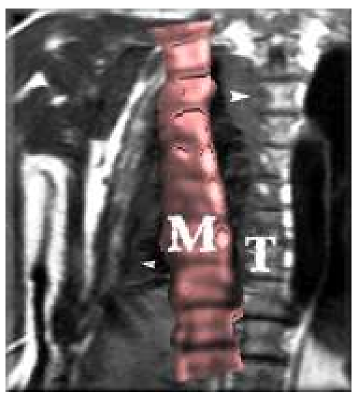

Figure 1: Model and target

The following segment expresses in more detail the full process involved. It comprises:

1. Appearance model construction.

2. Correlations learning in active appearance model.

[1]In the case of active shape models, on the contrary, inspection of nearby structures guided the model so that it better matched the target. Since the model landmarks usually reflect on the location of strong edges, similarity could be well-approximated by the distance between model points and strongest edges in their vicinity. Search along normals to the lines joining these landmarks led to appropriate edges location, provided that initialisation placed the model close enough to the target.

[2]This reverse process is used, for example, in application where reconstruction of faces or generation of flexible faces is of some value. Real-time animation is another possible extension of this.

3. Active appearance model search, based on resolved correlations.

The first and the second are concerned with training and learning whereas the third makes use of these previous two. It is somehow possible to learn from experience while searching, but it is not worthy of any further discussion.

Appearance Model Construction

The parameters which statistically describe the *shape* (much as in active shape models) can be expressed as a vector x, where

$$x = x_{mean} + P_s b_s. \quad (1)$$

x_{mean} (or $\overline{x}$) is the mean shape, as was calculated from the training data using, for example, Procrustes analysis[3]. P represents the eigenvectors of the covariance matrix (set of orthogonal modes of variation) and the parameters b_s control the variation of the shape. For n modes of variation, $1 < s < n$ holds.

Similarly, a vector g is used to describe the intensity of given pixels as derived from axis-aligned input data that is stretched to encompass the whole shape and fit or overlap the original model dimensions. Usually warps are used to displace the control points until they match those of the mean shape and shape-normalised patches can be captured. Just as before, variation is subjective to

$$g = g_{mean} + P_g b_g. \quad (2)$$

For shape, training is affected especially by the choice of landmarks identified in the image, whereas to extract intensity values a different approach is in use. This approach relies on the fact that geodesic interpolation can be applied to compensate for the noncontinuous results of the triangulation algorithm used. The linear form of the model as expressed above (1)(2) is due to Principal Component Analysis (PCA) which reduces the length of the vectors describing shape and texture, namely x and g respectively.

It is now imperative that the two equations above are merged in some way to create a new model that captures *both* shape and intensity. To do so, $b_s$and $b_g$are aggregated so they can be expressed as one single column vector

[3]Procrustes analysis has proved to be a popular method of shape analysis. The generalized Procrustes procedure was developed by Gower (1975) and has been adapted for shape analysis by Goodall (1991).

$$\left\{ \begin{array}{c} b_s \\ b_g \end{array} \right\}. \tag{3}$$

Applying further PCA, the following model is obtained:

$$\begin{array}{c} x_i = \bar{x} + Q_s c_i \\ g_i = \bar{g} + Q_g c_i \end{array} . \tag{4}$$

It is purely controlled by $c_1, c_2, ..., c_n$ where n is intended to be smaller than the number of b_s and b_g combined. That is simply due to the dimensionality reduction of PCA. Usually an inclusion of some weighing W is included to account for the difference in intensity value representation and the spatial cooridinates. The aggregation in such a case would take the form

$$\left\{ \begin{array}{c} Wb_s \\ b_g \end{array} \right\}. \tag{5}$$

but this is a practical consideration that need not be a concern at this point.

Learning Correlations in the Active Appearance Model

To improve the search performance, good choice of parameters adjustments is required. It is desirable to learn some correlations off-line and use them along with the model above to form a robust and efficient search. Observations are made to learn the correlation between the change in parameter values (usually each mode independently considered) and the pixel intensity difference that incurs. This means that for each change in the parameter values or for a collective change of several parameters, some change, in certain parts of the image in particular, will be quite apparent. A matrix of pixels (where rows represent horizontal scan lines in the image) is used to record the difference that a re-parameterisation imposes. More mathematically, for c_i which are the parameters as described above, a change δc is applied and the difference in intensities is calculated as follows:

$$\delta I = I_{model} - I_{image}. \tag{6}$$

Usually sum-of-squares is used here to penalise more harshly for blunt differences and ensure a summation of only positive values ($\forall x \subset \mathbb{Z} x^2 > 0$). Taking this intensity difference into consideration, the main correlation can now be expressed as:

$$c_i \rightarrow c_i + \delta c \rightarrow \delta I \tag{7}$$

which simply means that certain offsets to the parameters $c_i$cause a certain change in intensity. This correlation is recorded as follows:

$$\delta c = A \delta I \tag{8}$$

where A is a matrix recording the change in intensities due to the reparameterisation δc.

For each mode of variation and each pixel in the mean shape, weighting (negative or positive) is assigned to guide what the search will attempt to focus on. These "maps" of weights consume considerable amount of space, but are the only known paradigm for speeding-up through off-line computation. Wavelet compression can be used to reduce the space requirements and make active appearance models rather compact.

Active appearance model search Finally, using the model above and the correlation recorded for that model, it is possible to carry out the search as introduced in the beginning of this section.

The search is basically reliant on error or similarity measures calculated after each attempted parameterisation. Each such reparameterisation is initially guided by the matrices (images) that express correlating between the modes of appearance change and the intensity values.

The model, as shown in Figure 1, is placed within the image frame, close enough to its target. How close it should be put to the target is an issue that will not to be explained in any real detail, but true convergence may never be reached if bad initialisation takes place. The algorithm will most possibly then terminate when it reaches a local minima.

The basic search algorithm, expressed in a simplified way, is as follows:

- Take appearance model in its current state and do:
  - Compute δI, i.e. the difference between the model and the image.
  - Re-adjust parameter values c_i. Use the matrices learned off-line to make good choices.
  - Computer δI and save new appearance model if better results are obtained. If not, adjustment according to a coefficient $k = 1.5, 0.5, 0.25$ might be worth assigning to δc_i in order to achieve better results.
- While not converged or improvements are still being made.

The technique of matching an appearance model to an image is described in greater detail with some examples in [10]. It is also worth mentioning that in practice, in order to decrease the total run-time, varying increasing image resolutions are selected in the search iterations. This technique is a very common one in computer graphics and image interpretation tasks. A pyramid can be used to describe the data available for choice. The figure below shows how the size of the image quadruples (doubles in each of the two axes) at each stage of the pyramid, where the base of this pyramid is level 0 (level numbers increase upwards). Finer resolution images are at the bottom and low-resolution coarse ones at the top. The searching process typically begins at the very top of the pyramid and declares convergence only once it has reached the full resolution that is not lossy, i.e. it captures the whole pre-existent data.

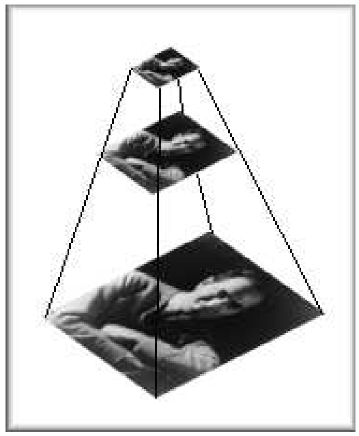

Figure 2: Resolution selection

Applications of Active Appearance Models

The strengths of active appearance models have often been demonstrated using facial images. Faces are a challenging type of data to cope with due to wide variation in gesture, inter-subject differences, facial hair, gender, lighting conditions and age which are some of the more dominant factors. It means that a good solution to face interpretation and recognition can lead to progress in other aspects of analysis, e.g. industrial inspection, medical data analysis, etc.

Commercial interest in this area has been another motive for increased experimentation with faces. Access control and gesture recognition are amongst the many possible uses of systems that investigate images (or sequences of images) of human appearance. The innate characteristics of the large number of different faces to be considered makes this problem far from trivial and it remains an excellent evaluation tool in the field of computer vision and not only model-based image analysis. Another frequently-used benchmarking data type is handwritten digits.

One common use of AAM's is for medical image analysis. In order to visualise shape and perform some measurements there needs to be a flexible model that handles anatomical variability and change, for example the expansion of an organ after some period of time. Contrariwise, some objects are not expected to show great variability, but abnormalities need to be detected. Industrial inspection products are an example of systems domain that follows such guidelines and they can often rely on edge detection and point-to-point measurements. These are simpler in the sense that they are quicker in operation and they resemble mammalian vision. Human eyesight is said to be highly sensitive to edges (c/f Mach banding). Sonka *et al.* [11] provide many more details on those topics.

2.2 Warps Representation

The potential of non-rigid registration of images has been a subject of research due to its ability to simplify the correspondence problem, amongst some other advantages it offers. As opposed to translation, scaling and rotation, all of which are rigid transformations, affine and non-rigid transformations [12] cater for flexible manipulation of points of interest. Folding and tearing has been the main drawback algebraic implementations of these transformation, but recently an interesting group of warps has been investigated. So-called *diffeomorphic* warps offer a solution to this drawback and they can easily be extended to 3-D in their regular form as shown in the figure below. Nevertheless, for most practical uses, a large number of such warps is needed, resulting in high computational demand. For further discussion of the application of non-rigid registration to landmark selection, see the work described in Hill *et al.*[13] and Rueckert *et al.* [14].

Figure 3 illustrates the effects current type of warps have on the space used to embed images. These warps are reminiscent of the ones described in Lötjönen and Mäkelä [15], but unlike many others, they have continuous derivates at the borders, which is a crucial condition for diffeomorphism.

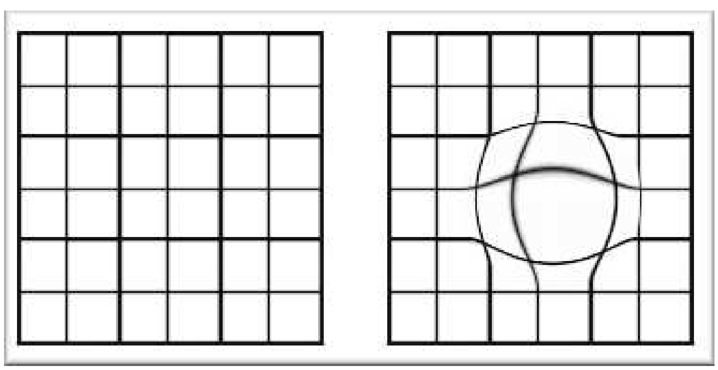

Figure 3: Warp example

When dealing with the aforementioned appearance models, an alternative emerges which chooses to deal with similarity measures using warps that minimise the difference between two images. Current research work attempts to apply the same principles to a large group of images and the result is a parameterisation that is compact in a global context. It relies on the many warps applied to the input data which bring their collective descriptive parameters closer together. As the different images are embedded in the heavily warped space, the spatial differences amongst the images are essentially being minimised.

2.3 Minimum Description Length

Overview

Minimum Description Length (MDL) can be used to drive the correspondence selection process so that better models are gradually produced. It is used to form an objective function that guides a minimisation aiming to find a good similarity measure, i.e. minimum apparent difference amongst a set, thereby choosing good correspondences. It is yet unclear how it can be usefully applies for appearance models. Figure 5 which is shown later illustrates the contribution of MDL to the overall process. It inputs data that is jointly generated from the current model and some data set and it outputs an estimate that adjusts impending warps and affects the choice of numerous parameter values.

MDL was extensively used in [4] where for some given model, a message is passed which gets assigned a value of length that implies complexity. The message is encoded to encapsulate the relation between a data examples and the up-to-date appearance model. Sometimes an evaluation data example is based on the leave-one-out validation technique, meaning that a large number of examples will generate a statistical model and one will be used to evaluate the model[4]. If the model is too complex or not suitable to represent the data, the message that is passed will be greater in length.

MDL in Action

The following chooses to focus on the incorporation of MDL in shape models. The process aims to choose good landmarks along some curve without any human intervention.

At the start, correspondences across a set of examples are places quite arbitrarily. Usually, a path-length spread of the points is a sensible enough choice which means landmarks are equally-spaced. This allows maximum freedom of movements for all landmarks mutually. A model is then created for the whole set and its parameters $b_i$which can solely characterise it are used to evaluate its compactness. For good choice of correspondences across the set we expect *low* values of $b_i$as well as ones with low variance, that is, a small range of acceptable values. The correspondences are then shifted along the contour of the shape iteratively in a process known as *continuous reparameterisation*. It is not just mathematically continuous (as it is functionally defined so that it fits any scale), but it also potentially affects all

[4]The model is constructed using only $n-1$ examples where n is the total number of examples and is then evaluated in accordance with the single left-out example. This procedure is usually performed n times, with a different example put aside at each iteration and averaging is then used to estimate the relevancy of the model.

landmarks along the curve in a ripple-like/cascading behaviour. The reparameterisation is kept diffeomorphic so that no landmarks move beyond the position of their successor (or more generally, one of their two neighbours), a step that could result in tearing and/or folding[5].

Experience has discovered that examples within the training sets should be dealt with one at the time[6], evaluating the *whole* set and the model at each stage. Reparameterisation is usually defined by some transformation rules that are vital to get good and fast results. The reparameterisation is usually applied to a number of adjacent landmarks at a time and different scales are chosen at random as well as the location being affected. At the later stages of the reparameterisation process, it is usually expected that no real improvements will be made for large scale alternation attempts and these will therefore be discarded fully in favour of small scale alternations that make the final fine adjustments. Experience has also shown that ultimately good choice of landmark can be made *automatic* mainly due to the ability of evaluating the model from information theoretic point-of-view, namely MDL.

MDL is well described by Rissanen in [16, 17] and the world wide web at: `http://www.mdl-research.org/`.

2.4 Bags of Pixels

A recently published technique is said to be capable of finding good dense correspondence. It is described by Jebara in [18]. Images are said to be better represented as sets of vectors for this specific purpose, as opposed to vectorisation where fixed ordering is imposed by *concatenation* of the vectors. Pixels are represented by the common (X, Y, I) tuple and the ordering of these tuples is arbitrary (they are said to analogically be placed in a bag so an alternative notion would be *sets of pixel*). Ways exist in which good configurations for ordering these pixels can be found. This implies that vectorisation of the pixels is not the sole option for effective image representation. As the process of pixel ordering takes place, dimensionality reduction is indirectly performed which transforms the image into a volumetrically minimal subspace and this reduction outperforms principal component analysis by orders of magnitude. This is one of the points that make this idea so appealing, but it is still extremely slow[7].

The figure below pictures the difference between common approach of pixel ordering versus the alternative bag of pixels.

---

[5]For closed curves, convex in particular, such problems are minute yet not negligible. A *monotonically increasing* reparameterisation function will ensure that points along the curve will at no stage overlap or conflict with one another in some way. It also has the advantage of allowing any number of landmarks to be considered, so resolution constrained could be specified beforehand.

[6]There has been the temptation of optimising several examples from the set in a distributed manner.

[7]The algorithms currently used for demonstration purposes take 3 days to run, but substantial speed-up is expected soon.

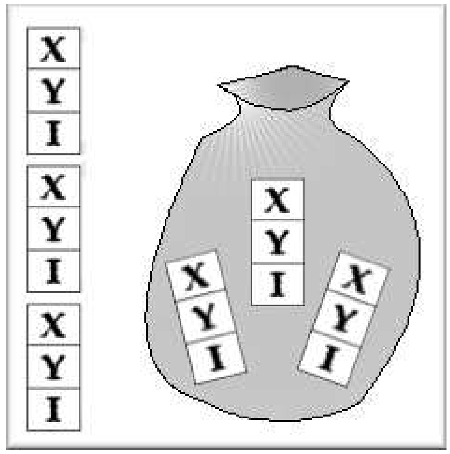

Figure 4: Bag of Pixels

3 Active Appearance Models Optimisation

3.1 Previous work

Minimum description length has been used previously to optimise models of shape and a similar approach can be used in part to improve the results of AAM formation. However, no analogous consideration has been applied to intensity values or textures; modes of shape variation and modes of brightness variation prove to be quite distinct. The former has not been combined in a particularly helpful way with the latter and an artificial mixture of the two, which is explicitly controlled by a coefficient λ, has been the only implementation attempted.

As was described earlier, in order to generate good models of shape, or more specifically good identification of landmark, warps were applied so that commonalities across the whole set of images get accentuated and a new morphed grid holds all the data.

3.2 Current work

Some research and experimentation have been carried out with the intention of achieving good appearance models using warps and group-wise[8] optimisation techniques. Algorithms have been utilised which are able to manipulate one item of visual data to match another using diffeomorphic warps (round in 2-D or spherical in 3-D). Such process must encompass

[8]Another advancement in current research is the inclusion of the whole set of data rather than just a one-to-one (pairwise) correlation between data and the current AAM. An investigation of just a couple at a time leads to poorer results in later practical use. That are due to the limited scope of the approach.

a large set of data in order to reliably generate a good active appearance model. Cross-validation is used to repetitively evaluate whether the model is altered appreciably or not. If a given warp appears to have the opposite effect, namely increase the difference between the model and the target or even leave it unchanged, it should be then discarded and the search for effective combination of warps virtually backtracks.

Some aspects of this current research are heavily based on the work of Davies *et al.*, but intensity attributes of the data are quite badly handled, especially given the high-performance of other components in the whole active appearance model. No existent evidence indicates that the current solution is the best solution or even a good one; in fact, quite the contrary holds. Better knowledge of the variation of intensity and its dependency on shape needs to be acquired first. It is still unknown whether any real correlation as such exists and, if so, which approach can capture it faithfully. Current approaches base this correlation on experimental evidence. In other words, textures are extracted from the training examples and recorded as a vector which is also statistically reliant on shape.

Another issue that is to some extent open for discussion is the procedural approach of geometrically transforming an image or image space. The methodology, precedence and ordering in which warps should be applied are not obvious. There is some satisfactory evidence though that current work surpasses its predecessors. Diffeomorphic warps and the issues related to them remain beyond the remit of the upcoming research work, yet it is crucial that their properties are fully realised. Since they affect both shape and texture, they have an impact on later stages of AAM optimisation.

Figure 5 can finally be presented as many of its constituent parts have been elaborated on. The images at the top of the figure are not all required to fulfil some predefined conditions (for example, having a large rectangular region at the centre), but some similarity between them is essential if valuable outcomes are sought. The warps applied to these images make group-wise registration possible. Comparison of each image to the reference image is still a valid choice, but empirical evidence suggests that results will then be inferior.

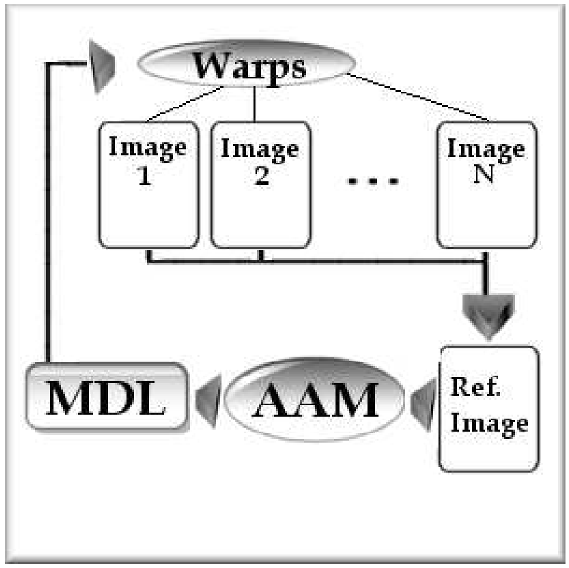

Figure 5: Active Appearance Models optimisation

3.3 Alternatives

From the graphical description above it can be somehow assumed that pipelining is worth investigation so that each of the modules in the cycle is kept busy. This of course would require an architecture that supports parallelism and is therefore not quite usable in typical work environments. Also, due to really heavy dependency in this cyclic system, many guesses need to be made and effort is thus being going wasted.

Some specific points of weakness were often described in technical reports in the field. There are many alternatives that were not taken into consideration, often because they did not appear trivial or had not been used before. It is often concluded that such alternatives are not constructive, whereas more widespread algorithms show higher success rates. For instance, no apparent attempt has been made to assign dynamically changing weights to the different components in the objective function and no wide range of warps has been made available in the transformation "arsenal".

4 Potential Development

There is no question about the possibility of improving previous results. Nonetheless, what needs to be pursued are ways of using *obvious evidential* methods to find models that are essentially *optimal* in some sense and are independent of the data under consideration. Algorithm that are *ad-hoc* and work better under certain conditions are of little interest as they will not generalise or offer any substantial progress in the long run.

Currently, the extensively used warps are not as flexible as one could hope for. While the theoretic principles work adequately, they sometimes prove impractical for use on contemporary machines. Better ways of applying such warps are being investigated so that fewer warps of lower complexity end up bringing two data samples to convergence rather quickly.

Texture patches have not seen significant enhancements in some previous work and further exploration seems worthwhile. Some of the principles from the work on ASM's could possibly be incorporated to produce more integral and consistent appearance models. The new models should exhibit good correlating between shape and intensity.

Some of the suggestions above are rather hard to convert into concrete implementations. Any such developments could lead to a real breakthrough. Experience and understanding of the research domain could suggest that a process of gradual trial-and-error would be fruitful. The next section presents some of the expected difficulties and Section 6 concludes and predicts the work plan that this document will entail.

5 Challenging Issues

Some of the more interesting issues are to do with feasibility. The mathematics behind diffeomorphism is said to be "not fully understood" and its application to computer vision is unprecedented[9]. Ensuring that any existing techniques and experiments remain valid in a space of high number of dimensions is another area that is hard to reason about and usually involves some trade-offs and simplifications. Several techniques that work perfectly well in 2-D can be completely useless in 3-D.

What makes this work slightly less worrisome is the proposition of new and better ways of achieving good models of appearance and low error rates, as described in previous papers with similar aims. These proposed steps are usually meaningless, however, if they cannot be backed-up by some ground-truth or a mathematical proof. Fiddling about with parameters and using *a-prior* knowledge of the problem is a logically good approach, but it contributes very little towards genuine research and exceptional insights.

6 Discussion

6.1 Summary

We presented some of the main concepts that should be useful in understanding the existing problems and improving past results. Many of these

[9]Comment: I am making some risky guesses here (as I often did beforehand). **Apologies for any disturbing assumptions that I make to preserve good flow.**

concepts already form the basis for analysis of related past work the experiments carried out throughout this work. Some of the weaknesses of previous work have been identified and advancements or alternatives have been suggested. The suggestions listed are slightly ill-considered in the sense that they have not been fully thought through and whether they will work or not is a question that will be answered as the project moves ahead. Real alternatives will be revealed once the work reaches its set milestones and is considered successful. Many more challenging issues are sure to come up and the work will most possibly have aims that cannot necessarily be achieved. Any contradiction to this assumption will make this work valuable.

6.2 Future Work

It is preferable to focus on one single aspect of optimisation, although several aspects can indirectly come together and overlap one another. Once this aspect or aspects are fully realised, they will occupy the whole work effort. If one path in the overall work fails or seems to have limited capacity for improvement, then effort will be diverted elsewhere with the intention of increasing research productivity. Every piece of results shown following this work will also be adjoined by the observation that some aspects are not worthy of further exploration. Consequently, good guidance on continuation of this research should be available.

The main direction that this work will take is yet unknown and it is therefore hard to say anything about feasibility considerations. The actual work to be embarked on and the objectives and milestones that go with it are due to be determined.

6.3 A Vision

From a system that has a collection of conjectured points of interest, a crude combination of two aspects of data collection and warping that is based on a vague emerging technique, we may soon be able to integrate some of the existing algorithms to form a powerful descriptor of appearance that can *automatically* learn about the mixture of elements in images and produce models that can be confidently labelled *optimal* according to some criteria. This system will be more efficient and more responsive, regardless of the type of data being inspected and its results will make it the best choice off-the-shelf.

Acknowledgements. The author would like to thank Chris Taylor for lengthy helpful discussions. This project is being funded by the EPSRC.

References

[1] M. Mitchell Tom, Machine Learning. New York ; London: McGraw-Hill, 1997.

[2] C. J. Beeston and C. J. Taylor, "Automatic landmarking of cortical sulci," in Medical Image Computing and Computer-Assisted Intervention - Miccai 2000, vol. 1935, Lecture Notes in Computer Science, 2000, pp. 125-133.

[3] A. D. Brett and C. J. Taylor, "A framework for automated landmark generation for automated 3D statistical model construction," in Information Processing in Medical Imaging, Proceedings, vol. 1613, Lecture Notes in Computer Science, 1999, pp. 376-381.

[4] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "A minimum description length approach to statistical shape modeling," IEEE Transactions on Medical Imaging, vol. 21, pp. 525-537, 2002.

[5] R. H. Davies, C. J. Twining, P. D. Allen, T. F. Cootes, and C. J. Taylor, "Shape discrimination in the hippocampus using an MDL model," in Information Processing in Medical Imaging, Proceedings, vol. 2732, Lecture Notes in Computer Science, 2003, pp. 38-50.

[6] R. H. Davies, C. J. Twining, T. F. Cootes, J. C. Waterton, and C. J. Taylor, "3D statistical shape models using direct optimisation of description length," in Computer Vision - ECCV 2002 Pt Iii, vol. 2352, Lecture Notes in Computer Science, 2002, pp. 3-20.

[7] T. F. Cootes, G. J. Edwards, and C. J. Taylor, "Active appearance models," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, pp. 681-685, 2001.

[8] G. J. Edwards, T. F. Cootes, C. J. Taylor, "Advances in Active Appearance Models," Seventh IEEE International Conference on Computer Vision, vol. 1, pp. 137-142, Kerkyra, Greece, 1999.

[9] G. J. Edwards, C. J. Taylor and T. F. Cootes, "Interpreting Face Images Using Active Appearance Models," Third IEEE International Conference on Automatic Face and Gesture Recognition, pp. 300-305, Nara, Japan, 1998.

[10] T. F. Cootes, C. Beeston, G. J. Edwards, and C. J. Taylor, "A unified framework for atlas matching using Active Appearance Models," in Information Processing in Medical Imaging, Proceedings, vol. 1613, Lecture Notes in Computer Science, 1999, pp. 322-333.

[11] M. Sonka, V. Hlavac, and R. Boyle, Image processing, analysis and machine vision. Pacific Grove, Calif. ; London: PWS Publishing, 1999.

[12] J. V. Hajnal, D. L. G. Hill, and D. J. Hawkes, Medical image registration. Boca Raton, Fla. ; London: CRC Press, 2001.

[13] A. Hill, C. J. Taylor, and A. D. Brett, "A framework for automatic landmark identification using a new method of nonrigid correspondence," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, pp. 241-251, 2000.

[14] D. Rueckert, A. F. Frangi, and J. A. Schnabel, "Automatic Construction of 3D Statistical Deformation Models Using Non-rigid Registration," in MICCAI, pp. 77-84, 2001.

[15] J. Lötjönen and T. Mäkelä., "Elastic matching using a deformation sphere," in MICCAI, pp. 541-548, 2001.

[16] J. R. Rissanen, "Stochastic Complexity in Statistical Inquiry" in World Scientific Series in Computer Science, vol. 15, Singapore, 1989.

[17] J. R. Rissanen, "Fisher Information and Stochastic Complexity", IEEE Transactions on Information Theory, vol. 42, no. 1, pp. 40-47, 1996.

[18] T. Jebara, "Images as Bags of Pixels," presented at ICCV03, Nice, France, 2003.

[19] A. Caunce and C. J. Taylor, "Using local geometry to build 3D sulcal models," in Information Processing in Medical Imaging, Proceedings, vol. 1613, Lecture Notes in Computer Science, 1999, pp. 196-209.

[20] A. Caunce and C. J. Taylor, "Building 3D sulcal models using local geometry," Medical Image Analysis, vol. 5, pp. 69-80, 2001.

[21] B. Moghaddam, C. Nastar and A. Pentland, "A Bayesian similarity measure for deformable image matching Source", in Image And Vision Computing 2001, vol. 19, issue 5, pp. 235-244.

[22] M. B. Stegmann, B. K. Ersboll, and R. Larsen, "FAME - A flexible appearance modeling environment," IEEE Transactions on Medical Imaging, vol. 22, pp. 1319-1331, 2003.

[23] K. N. Walker, T. F. Cootes, and C. J. Taylor, "Determining correspondences for statistical models of appearance," in Computer Vision - ECCV 2000, Pt I, Proceedings, vol. 1842, Lecture Notes in Computer Science, 2000, pp. 829-843.