ACCURATE MULTI-MODAL IMAGE REGISTRATION USING COMPRESSION

Pere-Pau Vázquez[1] and Jordi Marco[2]
[1]*MOVING Group, Universitat Politècnica de Catalunya - C/ Jordi Girona, 1-3, E-08034, Barcelona, Spain*
[2]*Dep. LSI, Universitat Politècnica de Catalunya - C/ Jordi Girona, 1-3, E-08034, Barcelona, Spain*

ABSTRACT

Image registration is an important task in medicine, especially when images have been acquired by different scanner/sensor types, since they provide information on different body structures (bones, muscles, vessels…). Several techniques have been proposed in the past, and among those, Normalized Mutual Information has been proven as successful in many cases. Normalized Compression Distance has been proposed as a simple yet effective technique for image registration. It is especially suitable for the case of CT-MRI registration. However, other image modalities such as PET pose some problems and do not achieve accurate registration. In this paper we analyse and propose a valid approach for image registration using compression that works properly for different combinations of CT, MRI and PET images.

KEYWORDS

Medical Image Registration, Image Compression, Visualization.

1. INTRODUCTION

Image registration is the process of overlaying two or more images that represent the same information taken at different times, from different viewpoints, and/or by different sensors. Aligning medical images has interest for the analysis of temporal patient evolution, the fusion of multimodal images, inter-patients comparison, and so on. We concentrate on multimodal images: images acquired with different processes, such as Computerized Tomography (CT), Positron Emission Tomography (PET), or Magnetic Resonance Imaging (MRI). Moreover, we will also consider only rigid transformations of the images. Our objective is to develop a method for robust multi modal image registration that achieves Mutual Information level results (one of the most prominent approaches in literature is Viola & Wells, 1997), requires little user intervention, and, if possible, is faster than previous methods. We use Normalized Compression Distance to solve the image registration problem. This measure has been previously used to address CT-MRI registration Bardera et al., 2010 based on the use of *bzip2*. Unfortunately, other image modalities, such as PET, yield poor results: registration point does not necessarily correspond to the correct one and, more importantly, the distance function has a high number of local minima, which will easily trap any optimization process. We improve this technique and achieve robust CT-MRI-PET registration by fundamental changes:

- Identification of the most suitable real-world compressor for image registration
- A new approach for image scan that provides better results than previous methods.
- Data quantization: Quantizing the input images allows us to improve registration quality and speed.

2. BACKGROUND

Normalized Compression Distance is a universal metric of distance between sequences. The **Kolmogorov complexity** ($K(x)$) of a string x is the length of the shortest binary program to compute x on a universal computer (such as a universal Turing Machine). Thus, $K(x)$ denotes the number of bits of information from which x can be computationally retrieved. As a consequence, strings presenting recurring patterns have low complexity, while random strings have a complexity that almost equals their own length. Hence, $K(x)$ is the lower-bound of what a real-world compressor can possibly achieve. The **conditional Kolmogorov complexity** $K(x|y)$ of x relative to y is the length of a shortest program to compute x if y is provided as an auxiliary input. Both Kolmogorov complexity and conditional Kolmogorov complexity are machine independent up to an additive constant. Bennett et al. (1998) define the **information distance** between two, not necessarily equal length binary strings as the length of the shortest program that, with input x computes y, and with input y computes x. The information distance is a metric, up to negligible violations of the metric inequalities. Li et al., 2004 present a normalized version of information distance, the *similarity metric*, defined as: $d(x,y) = max(K(y|x), K(x|y)) / max(K(x), K(y))$. The authors also prove that it is also a metric, and that this metric is universal: two files of whatever type similar with respect to a certain metric are also similar with respect to the similarity metric. Being Kolmogorov complexity not computable, it may be approximated with the use of a real-world compressor, leading to the Normalized Compression Distance (NCD): $NCD(x,y) = (C(xy)-min(C(x), C(y))) / max(C(x), C(y))$, where function $C(F)$ is the size of the compression of a certain file F, and xy is the concatenation of files x and y. Although the similarity metric has values in [0..1], NCD values are usually in the range of [0..1.1], due to compressor imperfections. NCD has been used for applications such as language classification and handwriting recognition in Cilibrasi & Vitanyi, 2005. They also analyse the conditions that compressors must fulfil in order to be used for computing the Normalized Compression Distance: i) **Idempotency:** For a repetition of a string, the compressor should be able to detect the repetitions and thus compress the file to a similar size than the original string compression $C(xx)=C(x)$, and $C(\lambda)=0$ where λ is the empty string. ii) **Monotonicity:** The concatenation of two strings should yield to a less compressible file than taking a single string alone, up to a certain precision: $C(xy)\geq C(x)$. iii) **Symmetry:** $C(xy)=C(yx)$. Compression should be symmetric, that is, changing the order of the concatenated strings should no affect the length of the compression. iv) **Distributivity:** $C(xy)+C(z)\leq C(xz)+C(yz)$. Real-world compressors seem to satisfy this property. Compressors with these properties are named *normal compressors*. Most real-world compressors fulfil those properties, at least to a point where they are usable for NCD computation. NCD has been used for music clustering (Cilibrasi et al., 2004), automatic construction of the phylogeny tree based on whole mitochondrial genomes (Li et al., 2001), the automatic construction of a language trees (Li et al., 2004). However, its use is not trivial, using NCD on the raw data may not yield good results, (Tran, 2007, Rocha et al., 2006). NCD has also been used for image classification (Cilibrasi & Vitanyi, 2005) with grey-scale images. Lan & Harvey, 2005 show that the measure performs better than histogram-based approaches in object recognition using PPM-based compression. A further work by Li & Zhu, 2006 improves the optimization task by Lempel-Ziv encoding and using either the dictionary, or the compressed patterns for measuring image similarity. Bardera et al., 2006 use Normalized Compression Distance for image registration. They select a *window* of pixels in one image and another one in the other reference image. Then, pixels are interleaved forming a new image where the red channel holds the pixels of reference image 1 and green channel the ones of reference image 2. These images are then compressed using JPEG 2000 and the compressed size is used as $C(xy)$ in NCD equation. They also present a second approach where the gray-scale values are treated as elements of a string, and *bzip2* is used to compress the resulting string. Again, the values of both images are interleaved. Although this approach works for the CT-MRI registration, it has problems for the PET-MRI image pair. The existing bibliography does not perform an in-depth study on the suitability of the different compressor families for any of the image comparison problems addressed. Image scan directions and concatenation building have also not being analysed, further than pointing out different possibilities (Bardera et al. (2006), Macedonas et al. (2008)). This paper intends to provide some answers to these questions for the concrete problem of medical image registration.

3. MULTIMODAL REGISTRATION USING COMPRESSION

3.1 Compression Algorithms

In general, a data compression algorithm focuses in identifying and extracting data redundancy. There are several ways to do this. In short, some of the main ideas under the standard compressors:

- **Huffman coding:** The core idea is to assign a fixed-length code to each symbol. The most frequent the symbol, the lower number of bits it is assigned.

- **Arithmetic coding:** The general idea is to replace a stream of input symbols with a single floating-point number. Compared to symbol coding, this method wastes a lower amount of bits.
- **Dictionary-based compression:** These methods encode variable-length strings of symbols as single tokens. These tokens are indices to a phrase dictionary. Thus, the compressor looks for redundant substrings.
- **Block-sorting compression:** It is based on a transformation that permutes the order of the characters. After the transformation, repeated characters are grouped together. This allows for other techniques that work on repeated characters (i.e. run-length encoding) to be applied to the transformed string.
- **Prediction by Partial Matching:** Adaptive statistical data compression technique that uses a set of previous symbols in the uncompressed stream to predict the next symbol in the stream (Cleary et al., 1995).

The main issue with multimodal image registration is that we are not looking for similarities but for *correlation*. This is due to the fact that images from different acquisition methods do not show the same information; on the contrary, it is somewhat complementary (bones versus soft tissues...). Therefore, the same grey level may indicate different information and be placed at different positions. That is why Mutual Information achieves so good results. Our objective is to identify the best compression scheme (if any) for image registration, and to determine which image sources combinations are feasible.

The Normalized Compression Distance works on strings. However, images contain 2D information that is not trivially changed to a 1D string (Macedonas et al., 2008). Thus, some possibilities arise, such as building a 1D sequence from an image by taking pixels row-by-row, or column-by-column. Macedonas et al., 2008 have tested both cases and stated that the row-by-row approach yields slightly better results. This is the approach also followed by Li & Zhu, 2006. Bardera et al., 2006 use *zig-zag* scan, and point as a future work the analysis of other 2D to 1D sequence conversion such as the use of space filling curves. In this paper we also analyse if the 2D to 1D transformation has any impact when using compressors for image registration. We will test different 2D to 1D sequence transformation methods: i) row-by-row, ii) column-by-column, iii) Space Filling Curves, and iv) random sampling.

We will see that random sampling improves image registration. Moreover, the *concatenated* file can be built by several different ways. The original formulation uses concatenation. However, Bardera et al., 2006 showed that pixel by pixel interleaving yielded good results. To the best of our knowledge, no previous study has addressed the effect of different ways to construct the *concatenated string*. Thus, we have evaluated several methods: i) Image concatenation, ii) Pixel interleaving, and iii) N-Pixel interleaving: the concatenated image is created by building groups of n pixels from each image. We used values of n=2 and n=4.

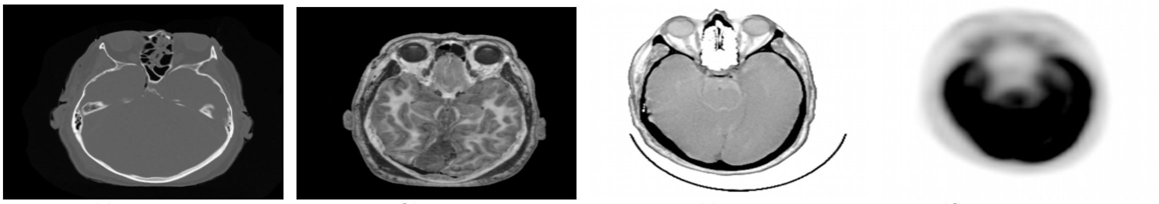
3.2 Experimental Setup

In this section we analyse the behaviour of different real world compressors for Normalized Compression Distance-based image registration using CT-MRI-PET source data. The experiments are performed by taking a pair of images that, otherwise indicated, are almost aligned. From those images, one is taken as the source, and the other one as the destination. In order to make the plots more intuitive, we only show transformations involving translations in both X and Y directions. For the tests, we move the source image over the destination one by applying translations from $(-10,-10)$ to $(10,10)$. Of course, the destination image has an extra frame of 10 pixels width (in background colour) around the original image, and we clip the destination image according to the position of the moving one. The results are shown as a 3D chart where we plot the Normalized Compression Distance using different compressors and parameterizations. Since the important point is the minimum value, and the function shape, due to the lack of space, we clipped the plots in order to show only the informative parts of the distance function. The minimum distance is the registration point. In order to compare the results, we will analyse the same image pair using Normalized Mutual Information. The desired result is a function that decreases as we approach the matching point and this one coincides with the one found with Normalized Mutual Information. Note that, for several reasons, the images coming from two different capture devices do not exactly sample the same regions in space. As a consequence, some small translation from the matching point given by NMI may be visually acceptable. A very important key issue is that the function should not be plagued with local minima, as this may challenge the optimization process.

Although we analysed a high number of compressors, we will only show the results from the ones that show better behaviour: a block-based compressor: *bzip2*, a compressor based on the Prediction by Partial Matching scheme: *paq8px*, one of the most effective data compressors, in terms of compression rate, according to Bergmans, 2011, but time consuming. Since some of the approaches are orthogonal, such as image scan versus image combination, throughout the paper we will analyse different configurations and incrementally incorporate them in the following sections, for the sake of the reading experience.

3.3 Image Concatenation versus Pixel-by-Pixel Interleaving

Bardera et al., 2006, 2010 showed that pixel-by-pixel interleaving was a good means to achieve CT-MRI registration with a block-based compressor. Despite that, the image combination possibilities were not deeply discussed. As a consequence, we first analyse the performance of different schemes using the same compressor. In contrast to the original approach, which analysed the image using *zig-zag*, we perform the image scan row-by-row, as we obtain equivalent results. An analysis on different image scans is presented next. The first experiment is CT-MRI registration. The images used are shown in Figure 1-left.



(a) CT image (b) MRI image (c) CT image (d) PET image

Figure 1. Left: CT and MRI pair to register. Right: CT-PET image pair.

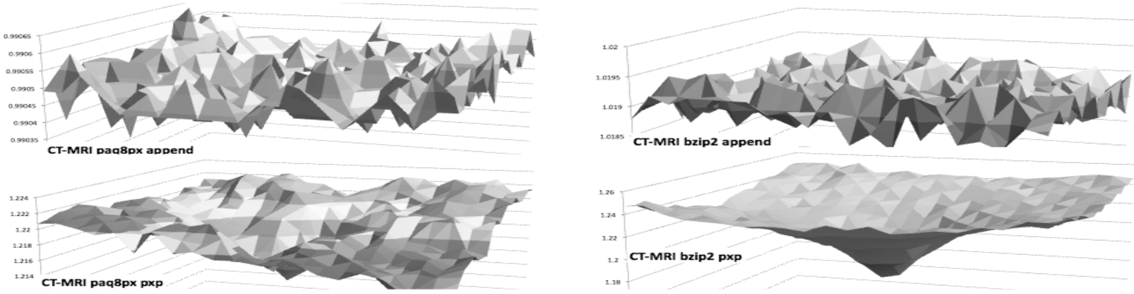


Figure 2. CT-MRI registration using different image combination methods: Regular file append (top for *paq8px* and *bzip2*, respectively) and pixel-by-pixel interleaving (bottom for *bzip2* and *paq8px*, respectively).

In Figure 2 we compare the different behaviour in registration when using image appending and image interleaving, respectively. Note that, independently on the quality of the matching point, pure image concatenation leads to a high number of local minima. From now on, the experiments shown will incorporate pixel-by-pixel interleaving.

3.4 2D to 1D Image Transformation

An image is a 2D data structure. For file writing, this information is transformed to a 1D array. This transformation can be done in different ways. So far, we have scanned the input images row-by-row. However, some other alternatives have been pointed out in literature. Since no experimental evidence on which method would yield better results for image registration, we have tested several approaches: row-by-row, column-by-column, space filling curves, and (pseudo-)random sampling.

Row-by-row and column-by-column have already been used for image comparison, with little advantage for row-by-row according to Macedonas et al., 2008. Bardera et al., 2006 suggested using Space Filling Curves. The rationale behind this is the fact that such curves take advantage of spatial coherence.

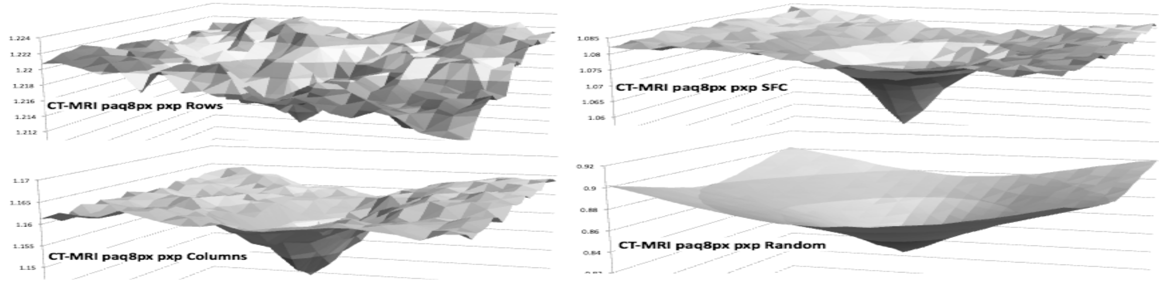


Figure 3. CT-MRI registration using different image scan strategies and *paq8px*. Column-by-column yields better results than row-by-row. However, the best matching is always found using random scan (bottom right): the matching point is good, and the function shape is soft.

From our experiments, we found that none of the previous approaches was optimal for image registration, as in most cases the function still contains a high number of local minima (Figure 3). Then, we came up with a totally different solution: random sampling the input images. We build the combined image by pixels selected from pseudo-random positions in the input images (the randomly selected position is the same for both input images). The reasoning behind is, when we are addressing image registration, we are not looking for local coherence, but pixel correlation between the two input images. Space Filling Curves may worsen the results because the information contents may vary greatly from one image to the other, and therefore, we are, somehow, counteracting the compressor task. For dictionary-based compressors, the dictionary construction might benefit from having a larger set of smaller words, than a smaller set of larger words. This is what may be induced by random image sampling. Actually, this technique improves image registration with dictionary-based compressors such as *paq8px* and *gzip,* and even block-based compressors such as *bzip2*.

We show the results of these different configurations in Figure 3 for the CT-MRI pair in Figure **1**-left. Note that random scan (bottom) produces higher quality results than with other techniques, both in terms of function shape, and a more accurate matching of the registration point. Row-by-row and column-by-column often obtain different, contrary results, but none is always better than the other. Sometimes row-by-row scan generates a softer shape, and sometimes column-by-column generates a better one. The results are unpredictable. We have not been able to find a consistent behaviour throughout the different tests. Moreover, except for the random sampling, none of the previous image scan methods improves the registration point.

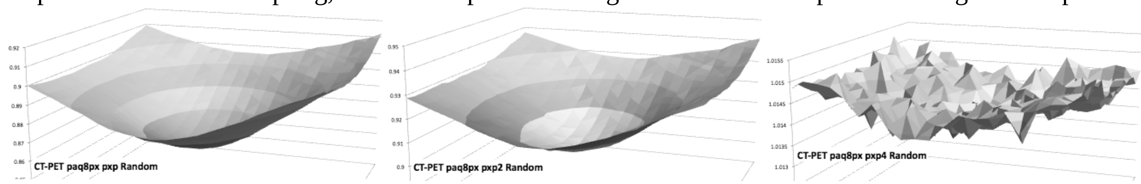
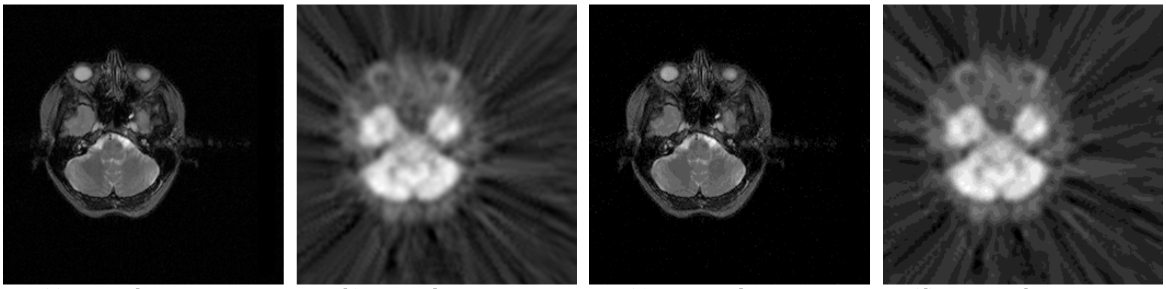


Figure 4. CT-PET image registration using *paq8px* and different ways to combine image pixels interleaving: single pixel interleaving (left), 2-pixel interleaving (center), and 4-pixel interleaving (right). Note how the best results are obtained with single pixel interleaving.

As already commented, we previously experimented with different interleaving strategies with no positive results. In order to assess our intuition supporting the random sampling strategy, we experimented again with N-pixel interleaving. If our intuition is truth, incrementing the number of pixels taken into account in the interleaving process should worsen the registration function. We used the CT-PET image pair in Figure **1**-right. This is shown in Figure **Erro! A origem da referência não foi encontrada.** where pixel interleaving sizes are compared: one (pxp), two (pxp2) and four (pxp4) pixels. The method that obtains better registration is always 1 pixel wide image combination. This enforces our idea that correlation is better captured if we randomly sample the input images.

3.5 Improving Registration using Image Quantization

When addressing image registration using Normalized Mutual Information, the number of bins selected for the histograms does influence the registration results. More concretely, the registration may improve if we select a lower number of bins, say 100 for instance. Furthermore, a lower number of bins also accelerates the registration algorithm because the joint histogram is sensitively simpler. This bin reduction also makes the algorithm more robust to the presence of noise. In our experiments we follow the same idea. We quantize the input images to 16 bins (as shown in Figure 5). This reduces the noise and the amount of information, but the important details are not removed. As a consequence, registration results are improved.



(a) Original MRI (b) Original PET (c) Quantized MRI (d) Quantized PET

Figure 5. Quantization of the input images to 16 values. Left column shows the original images to register, and the right column shows the same images after the quantization.

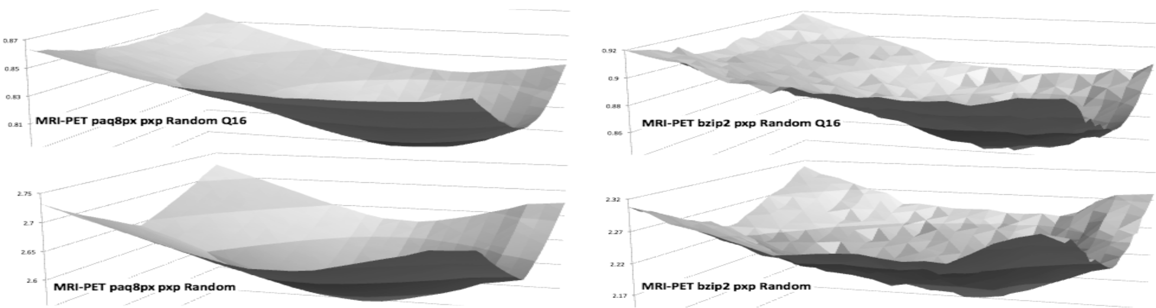


Figure 6. Comparison of MRI-PET registration using *paq8px* and *bzip2* using quantization and with unmodified image sources. Although quantization does not improve the function shape, it improves the matching point (with NCD and *paq8px* we match the result obtained by Normalized Mutual Information).

Once the images are quantized, we compare both using the Normalized Compression Distance. We set one of the two images as the source, and the other one as the destination. Again, we move the source image over the destination one by applying translations from $(-10, -10)$ to $(10,10)$. The minimum distance is the registration point. As shown in Figure 6, we are able to correctly register the MRI-PET pair, a difficult example, since the information in the PET image is not very detailed.

3.6 Results

We have tested our registration scheme with several multimodal image pairs, and some of them already appeared in this paper. So far, we have only analysed the adequacy of the shape of the distance function. In this section we will further analyse the obtained registration points with a pair of CT-MRI, another one consisting in CT-PET, and finally, we also test an MRI-PET pair.

From now on, the experiments are carried out using all improvements: image quantization, image interleaving, and random scan. For the sake of clarity, we plot the NCD values obtained by *paq8px* compressor, since it is the one that showed most robustness in the experiments. Over the plot, which is shown as a contour chart, we indicate the registration point obtained with this compressor. We also add the registration point obtained by Normalized Mutual Information and the results we obtain with our algorithm and other compressors such as *bzip2* and *7z* (PPM-based, with results similar to *paq8px* but much faster).

Our method performs **CT-MRI registration** very efficiently (with good results even for *gzip*). Figure 7 (left) shows the results for the images in Figure **1**. In this case, all the registration points computed by our method with different compressors (*paq8px*, *bzip2*, and *7z*) correspond to the value computed using Normalized Mutual Information. In this case, *gzip* also correctly finds the registration point.

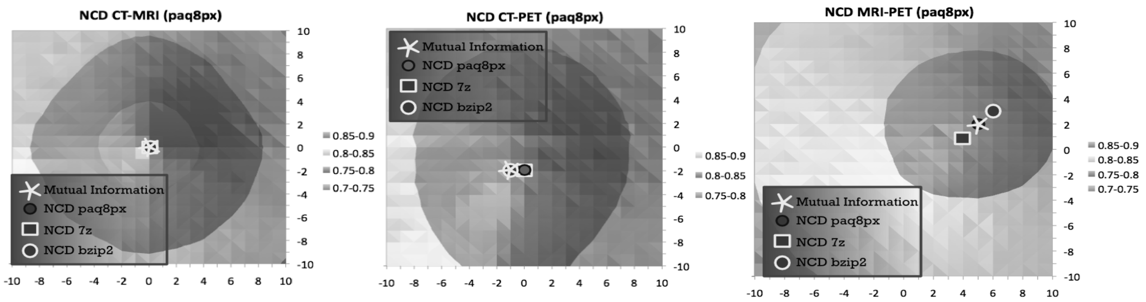


Figure 7. Different registration methods with different image modalities. Left: CT-MRI registration. All registration points coincide with the NMI method. Centre: CT-PET registration. NMI and *bzip2* achieve the same result, while using *paq8px* or *7z* obtains a point shifted only one pixel. Right: MRI-PET registration.

As already said, **CT-PET registration** is usually difficult because of the lack of details of PET images (see Figure **1**-right). If we do not use our improvements, *bzip2*, is unable to match the registration point. However, with pixel interleaving and random scan, results are highly improved. Figure 7 (centre) shows the results obtained with our algorithm. Here, block-based compression obtains the same result than NMI, while *paq8px* and *7z* are shifted a pixel in the *X* direction. This is not bad at all, since, as noted previously, the images do not exactly sample the same region of the body, and both matching points are visually acceptable.

The **MRI-PET registration** is also difficult due to the lack of details in PET images. However, our method correctly finds a registration point for these kinds of source data. We show the results in Figure 7 (right). Note that *paq8px* finds the correct point according to NMI, but both *bzip2* and *7z* achieve very good results. In all cases, the matching points are visually acceptable.

4. CONCLUSION

We presented a novel approach to multi-modal image registration using compression that obtains robust results even with images such as PET. Accurate registration is obtained by modifying the *classical* NCD measurement in three ways: a) We quantize the input images, b) we scan the images in a pseudo-random manner, and c) we combine the images pixel by pixel to form the *concatenated* file.

Throughout the process, we also made other interesting findings: a) for image scan, Space Filling Curves and row-by-row or column-by-column did not give good results. b) regular image concatenation is also unsuitable for image registration. c) PPM-based compressors are more robust for image registration than other schemes. Our results are comparable to registration by Normalized Mutual Information. Hence, we believe that PPM-based compressors are a promising tool for further investigation. Though *paq8px* is very costly, another PPM-based compressor, *7z* works very fast, and has also proven useful for image registration. Its results outperform NMI times even when computing NMI with a reduced number of bins such as 100.

We also tested other compressors, such as *jpeg2000*, *rzip* (tailored to find redundancies placed at high distances), *lzma* (based on Lempel-Ziv and Markov Chain coding), *hffzip* (Huffman coding), and *gzip* (based on Lempel-Ziv). Except for *gzip*, that gives good results for CT-MRI registration, the other ones had not satisfactory results in any case. In future we want to deal with image sources such as SPECT.

ACKNOWLEDGEMENT

This project has been supported by TIN2010-20590-C02-01 of the Spanish Ministry of Science and Innovation. We want to thank the images providers, especially the "Vanderbilt Database" and Pubimage sets. This paper was partially developed during a sabbatical period by the authors at the Geometry and Graphics Group from Universitat de Girona.

REFERENCES

Bardera, A. et al, 2006, Compression-based image registration. *IEEE International Conference on Information Theory*, pp 436–440.

Bardera, A. et al, 2010, Image registration by compression. *Information Sciences*, 180(7), 1121 – 1133.

Bennett, C. et al, 1998, Information distance. *IEEE Transactions on Information Theory*, 44.

Bergmans, W., 2011. Maximum compression. http://www.maximumcompression.com/. [Online; accessed Feb 2011].

Cilibrasi, R. and Vitanyi, P.M.B., 2004. Clustering by compression. *IEEE Trans. Information Theory*, 51(4):1523–1545.

Cilibrasi, R. et al, 2004. Algorithmic clustering of music based on string compression. *Journal of Computer Music*, 28(4):49–67.

Cleary, J.G. et al, 1995, Unbounded length contexts for ppm. *DCC '95: Proc. of the Conference on Data Compression.*

Lan, Y. and Harvey, R., 2005, Image classification using compression distance. *Proceedings of the 2nd International Conference on Vision, Video and Graphics*, pp 173–180.

Lawder, J.K. and King, P.J.H., 2000, Using space-filling curves for multi-dimensional indexing. *17th British National Conference on Databases*, pp 20–35, London, UK.

Li, M. et al, 2001. An information-based sequence distance and its application to the whole mitochondrial genome phylogeny. *Bioinformatics*, 17(2):149–154.

Li, M. et al, 2004, The similarity metric. *IEEE Transactions Informmation Theory*, 50(12):3250–3264, 2004.

Li, M. and Zhu, Y., 2006, Image classification via lz78 based string kernel: A comparative study. *PAKDD*, pp 704–712.

Macedonas, A. et al, 2008, Dictionary based color image retrieval. *J. Vis. Comun. Image Representation*, 19(7):464–470.

Rocha, J. et al, 2006, Compression ratios based on the universal similarity metric still yield protein distances far from cath distances. *CoRR*, abs/q-bio/0603007.

Tran, N., 2007, The normalized compression distance and image distinguishability. *Proceedings 19th IS&T/SPIE Symposium on Electronic Imaging Science and Technology*, San José, USA.

Viola, P. and Wells, W.M., 1997, Alignment by maximization of mutual information. *J. Comput. Vision*, 24(2):137–154.