Novel motion detection algorithm of 3D medical data for diagnosis purposes

Martin Zagar[1], Alan Mutka[1] and Hung-Cheng Tai[2]

[1] RIT Croatia, IT / WMC Department

Don Frana Bulica 6, 20000 Dubrovnik, Croatia

[2] Department of General Education, Chiayi Campus, Chang Gung University of Science and Technology, Chiayi 61363, Taiwan

ABSTRACT

The encoding and streaming of medical video imaging for diagnosis purposes is an example of high computational capacity applications. Motion detection presents one of the pillars in encoding such videos and a basis for compression where just motion vectors are encoding instead of the complete frame. We propose a new algorithm for motion detection of 3D medical videos where the novel approach is based on the fact that medical videos don't have sudden and large motions. We defined an adaptive medical search algorithm that starts with presumptions that the motion is coherent for acquired medical 3D MRI or CT data.

Keywords: Motion detection, 3D medical data, Data compression, Search algorithms

INTRODUCTION

Encoding and sending medical video data for diagnostic purposes is an example of a large computing application. Motion detection is one of the initial steps in encoding such videos and a basis for compression (Motta, 2000) where only motion vectors are encoded instead of a complete data frame with fewer data generated after encoding. This is very important because medical data can be quite demanding in terms of size, indicating problems with storing and sending large data sets over the Internet, as referenced in (Sriraam, 2008). The challenge is to define a framework for highly efficient compression for coding medical data in remote diagnostics by proposing a new motion detection algorithm in 3D medical videos.

The aim of this paper is to propose a flexible medical data search algorithm based on the assumption that the shift between the two image frames is coherent and minimal in the context of medical 3D MRI or CT data. This would assume an input hypothesis to improve performance over existing DICOM browsers (medical data visualization applications). This could improve the availability of standard healthcare to patients (given the fact that the amount of coded data would be reduced, existing DICOM browsers could run on smartphones and be available wherever there is a medical specialist, which is important in remote diagnosing and enabling a second opinion of medical experts). Also, system requirements would be reduced and data browsing speed in DICOM browsers would be increased, and storage requirements would also be reduced as our solution provides significant savings compared to the current standard solution in which the Full-search algorithm is applied

Currently in the H265/HEVC encoding standard (the most recent standard for video compression), in sense of accuracy, only the Full search algorithm for motion estimation of medical data analysis is applicable, so there is a need for implementation of other, more efficient algorithms. While motion estimation algorithms for general-purpose algorithms (where input video can have any possible motion features), as (Dikbas, 2010) and (Neves, 2009), are well known (triangle, pentagon, and square pattern), those for medical applications without sudden and large motions are not yet optimized which is very relevant in sense of speed needed for calculations, when we consider typical 500MB MRI datasets.

In this paper, we will present our research outcomes, main benefits, and possible risks in implementation, and also describe how our novel algorithm for motion detection of 3D medical videos is highly efficient while the same accuracy as the currently used full search algorithm for motion detection in 3D medical data videos.

METHODS AND MATERIALS

As noted, only the Full search algorithm is currently implemented as part of the H265 / HEVC coding standard, and our goal was to define an algorithm that would be more efficient given the above assumptions and that could be efficiently used in more complex applications such as (Klapan, 2019a) and (Žagar, 2021). As part of this, we analyzed motion estimation algorithms for 3D videos in two directions - in one direction are general-purpose algorithms where the input video can have all possible motion features, and the other is strictly for medical application. What we are proposing is a new framework for motion detection in 3D medical videos. The new approach is based on the fact that medical videos do not have sudden and large movements, so the first idea was to adapt some algorithms that predict small simple linear movements, such as the existing One-at-the-time (OTS) search from 2D to 3D. The advantage of this approach is that it is very effective, but can be limited to local minimums and therefore predict movement in the wrong way which is definitely undesirable for medical purposes. Therefore, we have defined a flexible medical search algorithm that starts with the assumption that the motion is coherent since the input data were obtained from an MRI or CT device.

Our algorithm is presenting an entry point of a bigger framework for the compression of multidimensional medical data. In these terms, a list of potential customers includes any users of medical data visualization systems that seek high accuracy and fast computation for multidimensional medical data visualization, for example, in applications in OR during surgeries, preoperative planning, and diagnostics, that might be used by a wide variety of medical specialists currently using regular DICOM viewers with inefficient rendering. We plan to offer our algorithm to producers of the DICOM viewers to optimize their calculations for visualizations, to speed them up.

Compared to the Full search algorithm, in our algorithm, the initial search pattern is checking just 23 points, one center point, 14 checking points that surround the center (at points with absolute Manhattan distance equals 2), and additional 8 points with the following spatial vectors

$$|\mathbf{i}|=|\mathbf{j}|=|\mathbf{k}|=1\ . \tag{1}$$

This is to obtain maximum accuracy for the initial motion vector whereas there is a tradeoff for further calculations. Framework next predicts motion vector for all other consecutive frames based on minimum block distortion for two possible blocks – the one which is predicted by the previous motion vector and the new one re-calculated with our algorithm.

Algorithm Design

In this section, we are presenting an algorithm proposal based on the above assumptions. It is important to note that the output algorithm provides a shift vector in a computationally efficient way compared to the full search algorithm. The initial step of our proposed algorithm is equal to the initial step of the Diamond Search (DS) algorithm and looks like this:

The initial step. The initial search pattern focuses on the origin of the search window, and the measure of interference between blocks is calculated using the SAD algorithm

(Sum of Absolute Differences) for 9 control points, one center point, and eight points surrounding the center, forming a diamond shape (at points where the absolute Manhattan distance is 2). This is shown in Figure 1. The initial threshold defining the measure of acceptable disorder is T = 10%.

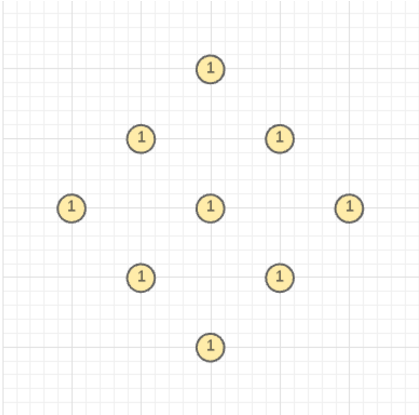

Figure 1. Initial search pattern

Step 1. If the point of minimum disturbance in the block (MBD) is more than the amount T less than the first following minimum and if it is also the center point:

- Find the minimum block disturbance at points with an absolute Manhattan distance equal to 0 or 1 from the current center. The identification of the current destination is shown in Figure 2 until the identification of points with an absolute Manhattan distance equal to 0 or 1 from the current center is shown in Figure 3.
- The point of minimal disturbance found in this step is the final solution of the motion vector pointing to the best matching block. This is shown in Figure 4.

Otherwise:

- Go to step 2.

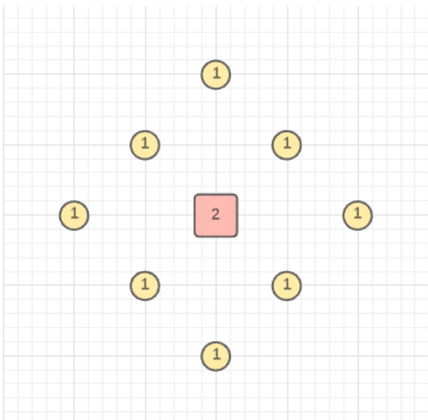

Figure 2. Identification of the current destination

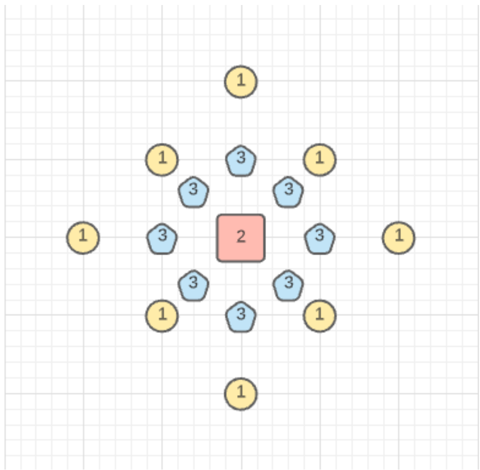

Figure 3. Identification of points with an absolute Manhattan distance equal to 0 or 1 from the current center

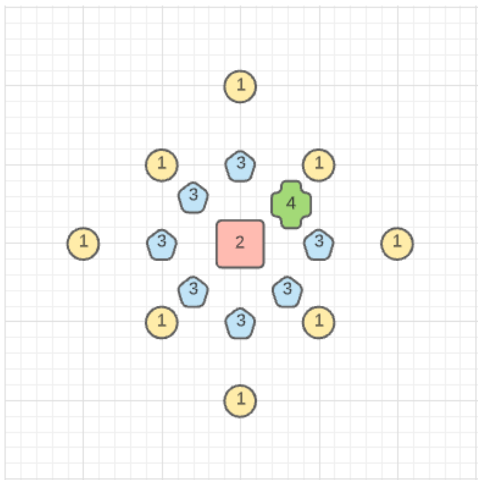

Figure 4. The case when the point of minimum disturbance in the block (MBD) is more than the amount of T less than the first following minimum and if it is also the center point; the algorithm ends

Step 2. If the point MBD is less than the amount T less than the first following minimum but is not at the position of the center point (the steps of the algorithm, in this case, are summarized in Figure 5):

- The block with the new MBD becomes the new focal point
- Find the MBD either at that point or at a point opposite the position of the previous center.

If the point MBD is less than the amount T less than the first following minimum):

- The first two minima are taken into account and the disturbance in the block is calculated for points located at opposite positions of the previous center with respect to these minima
- If (algorithm iteration number <3):
    - Go to step 1
- Otherwise:
    - The MBD point found in this step is finally the solution of the motion vector pointing to the best matching block
    - The threshold defining the measure of acceptable T disorder is increased by 5%.

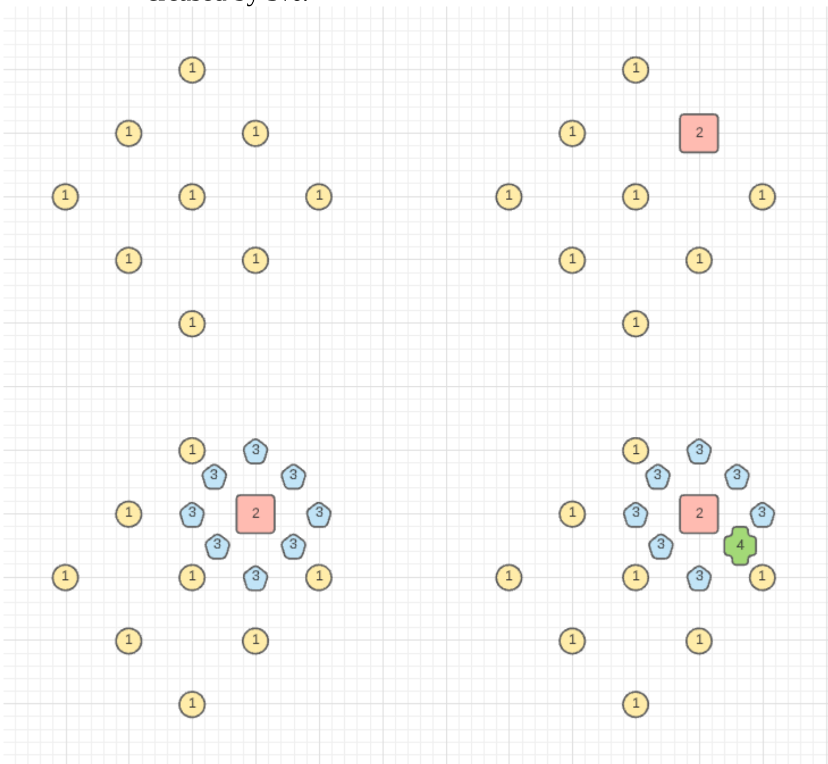

Figure 5. Summary of the algorithm for finding the displacement vector in case the point MBD is less than the amount T less than the first following minimum but is not in the position of the central

RESULTS AND CONCLUSIONS

If we compare our solution with the current standard motion detection solution for medical data compression, our solution will generate savings in the cost of storage space (the cost of our solution is 0.25 EUR compared to 10 EUR compared to the standard solution - the standard compressed DICOM study is 200 MB, using the proposed algorithm this size drops to 5 MB, with a cost of 1 EUR for 20 GB on iCloud), based on a one-year time frame in which typically medical professionals use data analysis and visualization of 1000 MRI / CT scans on one device.

Further savings will be achieved in access and bandwidth of the communication channel (the proposed solution would cost 15 EUR compared to the standard solution which costs up to 600 EUR, based on the channel bandwidth price of 3 EUR for 1 GB of mobile data); in system requirements - the proposed solution does not require complex and especially capacitive computer systems, but the algorithm could be processed on simple smartphones (usually for 300 EUR, compared to the typical price of 1000 EUR for a medium performance mobile computer). The total savings is close to 1300 EUR per year for visualization of 1000 MRI / CT scans on a single device (the cost of the solution using the proposed algorithm is around 310 EUR, and the standard cost is around 1600 EUR).

The main advantage would be reduced time and storage when running such applications which would increase the user experience (which could be important for medical professionals in the operating room), as referenced in (Klapan, 2019b), and reduce the cost of equipment needed to run such applications (less computational complexity will require fewer computers).

The main obstacle in the full implementation of the medical/diagnostic data visualization market solution is the need for a certificate for medically approved software, which is expensive and could take time. This risk of testing and certification is a major risk that needs to be considered before selling a product. We additionally identified a potential lack of market recognition as medical professionals are reluctant to change the software they use frequently.

There are currently no commercially available solutions with a range of functionalities similar to the proposed solution. All current DICOM browsers on the market use a non-optimized full search algorithm to detect motion in medical images. We see an opportunity in the context of updating image viewers for diagnosis purposes that are currently present on the market (like DICOM viewers), so the computations of motion detections will be much faster and equally accurate as when using the Full search algorithm. The main advantage would be time reduction when running such applications that would raise the user experience (which might be important for medical specialists in OR) and reduce the cost of the equipment needed to run such applications (because of less computational complexity it will need less computer power).

REFERENCES

Dikbas, S., Zhai, F.: Lossless image compression using adjustable fractional line-buffer, Sig. Proc.: Image Comm. 25(5): 345–351. (2010)

Klapan, I., et al.: Computer-assisted navigation, 3D-surgical planner, augmented reality in the operating room with 'in the air' surgeon's commands as "biomechanics" of the new era in personalized contactless hand-gesture non-invasive surgeon-computer interaction?, Biomed. J. Sci. Tech. Res. 19(5), 14678–14685. (2019a)

Klapan, I., et al.: Utilization of 3D medical imaging and touch-free navigation in endoscopic surgery: does our current technologic advancement represent the future innovative contactless noninvasive surgery in rhinology? What is next? Biomed. J. Sci. Tech. Res. 22(1), 16336– 16344. (2019b)

Motta, G., Storer, J.A., Carpentieri, B.: Lossless Image Coding via Adaptive Linear Prediction and Classification, Proceedings of The IEEE 88(11): 1790–1796. (2000)

Neves, A. J. R., Pinho, A. J.: Lossless compression of microarray images using image-dependent finite-context models, IEEE Trans. Med. Imaging 28(2): 194–201. (2009).

Sriraam, N., Eswaran, C.: An adaptive error modeling scheme for the lossless compression of eeg signals, IEEE Transactions on Information Technology in Biomedicine 12(5): 587–594. (2008)

Žagar, M., Mutka, A., Klapan, I., Majhen, Z.: Hand and Gesture Module for Enabling Contactless Surgery. Trends Telemed E-Health 3(1). TTEH. 3(1), 1-5. (2021)