Can a Human Operator Be Replaced By a Computer For Processing of Bolus-Tracking Perfusion Data?

M. Straka[1], M. Mlynash[2], G. Zaharchuk[1], G. W. Albers[2], and R. Bammer[1]

[1]Lucas Center, Department of Radiology, Stanford University, Stanford, CA, United States, [2]Stroke Center, Department of Neurology and Neurological Sciences, Stanford University, Stanford, CA, United States

Purpose: Measurements of cerebral perfusion deliver valuable information about the state of brain tissue in case of acute stroke and other cerebrovascular diseases. One of the available techniques for assessment of brain perfusion is bolus-tracking (DSC-MRI). Its use in clinical routine has been cumbersome because obtaining maps of perfusion parameters typically requires substantial human operator post-processing of the raw data, which is lengthy and subject to errors. To alleviate this limitation, we have been developing and implementing a system that aims to deliver perfusion maps in real time with no operator supervision. The key component of the system is an automatic selection of arterial input- (AIF) and venous output- (VOF) functions. The selection must be robust and reliable; otherwise the quality of computed perfusion maps suffers. Therefore a strong emphasis was put on achieving robustness in presence of noise and motion artifacts, as well as of unpredictable errors arising from varying acquisition parameters, T1-weighting in signals, and confounding effects of bulk-blood in vascular signals acquired with GRE-EPI sequences. The purpose of this study was to evaluate whether the automatic AIF/VOF selection algorithm can perform as well as or outperform a human operator in terms of speed, robustness and accuracy, and hence if the automated solution could overtake this tedious manual work.

Methods: The algorithm used to automatically select the AIF/VOF is a part of an in-house research software written in C++, aimed on computation of quantitative perfusion parameters (CBV, CBF, MTT, t_{max}). To reduce the influence of motion, which confounds many PWI scans, the system first executes motion correction in the image domain. Then, to select the AIF, an algorithm similar to [1] generally looks for signals that are: 1) in the anterior part of the brain, 2) have amplitude above average, 3) have the peak value earlier than average (but not earlier than estimated arrival of bolus), and 4) are narrower than average. To make the selection sufficiently robust, high importance is put to spatial clustering of the evaluated signals; the cluster size is a function of the data resolution. The signals in a particular cluster must be similar in timing parameters and amplitude. Any signals with distorted shape or excessive noise are rejected. Among the remaining clusters, multiple points with highest score are selected and the final AIF is obtained by averaging them. The VOF selection uses a similar approach for the posterior region, but includes an additional constraint requiring that the VOF peak should not appear earlier than 2s after and not later than 12s after the detected AIF peak. To quantitatively evaluate the above described algorithm, we used a set of DSC-MRI PWI from the DEFUSE study [2] (N=30, 10F/6M, age 32-92, 16 pre-treatment and 14 post-treatment scans) with different acquisition parameters (GRE-EPI, TR=1.44-2s, TE=41-60ms, FA=60°-90°, 9-15slices, slice thickness 5-7mm/gap 0-2mm). Within these data, 3 datasets contained significant amount of in-plane and through-plane patient motion. In the processing, first the software was run and the computerized estimates of AIF and VOF signals were obtained. Next, two clinical experts manually selected AIF and VOF in the same dataset, while the time needed to process the case manually was recorded. After the manual selection was done, the computerized results were shown to the experts and they were then asked to 1) compare their selection of the signals with the algorithmic result (using a 3 point scale: better, equivalent, worse) and 2) to decide if the computerized result was acceptable for PWI processing. The recorded manual and automated results were post-processed and resampled for statistical evaluation of the arrival time, FWHM, peak values and anatomic territories.

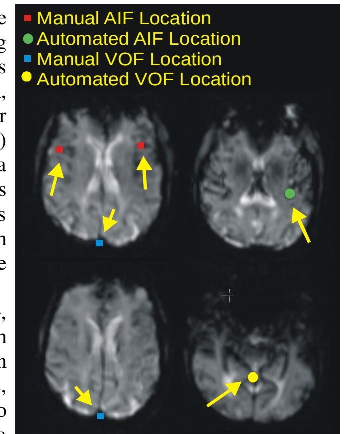

Fig. 1: Example of differences in AIF and VOF locations between manual reader and the algorithm

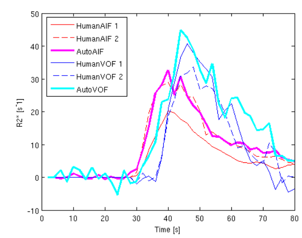

Fig.2 Example of the AIF and VOF signals selected by humans and the algorithm (in data from Fig. 1).

Results: The results are summarized in Tab.1 and Tab. 2, and are presented for first/second human reader. Only in 13%/30% AIF and 13%/17% VOF cases the shape/arrival time was considered worse then manual. More importantly, the readers agreed that in 100% of AIF cases and in 93%/100% VOF cases the algorithmic results were acceptable for further PWI post-processing, and only one reader considered 2 VOF cases (7%) as not acceptable. The typical time needed for human processing of the cases was 2:52min/3:23min median time (min 1:45/2:32, max 5:47/6:12). The computerized selection takes less than 5 seconds, using multi-threaded implementation on an Intel Pentium 2.0GHz CPU. The human readers consistently selected the AIF in middle cerebral artery (MCA), and VOF in the superior sagittal sinus (SSS), whereas the algorithm selected the AIF in MCA (N=22), anterior cerebral artery (N=6), internal carotid artery (N=1) and posterior cerebral artery (N=1). The automated VOF was selected in SSS (N=19), straight (N=9) and transverse sinuses (N=2).

	# of times (%) the signal shape was:			# of acceptable cases (%)	# of not acceptable results (%)	# of cases in which operator would change their decision (%)	# of cases where operator would rather keep his selection (%)
	Better	Equal	Worse				
Automatic vs. manual AIF	10 / 5 (33% / 17%)	16 / 16 (53% / 53%)	4 / 9 (13% / 30%)	**30 / 30 (100% / 100%)**	**0/0 (0% / 0%)**	9 / 5 (30% / 17 %)	6/9 (20% / 30 %)
Automatic vs. manual VOF	4 / 3 (13% /10%)	22 / 22 (73% / 73%)	4 / 5 (13% / 17%)	**28 / 30 (93%/100%)**	**2/0 (7% / 0%)**	3 / 3 (10% / 10 %)	5/2 (17% / 7 %)

Tab. 1: Ratings of automated selection of AIF/VOF by first/second human operator. The numbers in the left indicate the opinion of the human readers about the automatically selected AIF/VOF compared to their selection, in terms of shape and arrival time but regardless of the anatomic location. The two right columns summarize the willingness of the readers to trade or keep their selection, if also anatomic location was considered.

Comparing the selected signals (means for algorithmic results)	AIF arrival time [s]	AIF FWHM (mean: 10.1±0.1 [s])	AIF peak (mean: 45 ± 7[s^{-1}])	Area under AIF (mean: 960 ± 268)	Area under VOF (mean : 711±215)
between reader 1 and 2	0.76 ± 0.80	-1.32 ± 2.02	-3.02 ± 5.94	-148±212	-92±213
between reader 1 and program	0.80 ± 0.86	-0.54 ± 1.86	-0.96 ± 7.70	-88±173	-214 ± 256
between reader 2 and program	0.04 ± 0.82	0.78 ± 2.45	2.06 ± 6.70	60±214	-122±191

Tab. 2: The summary of parameter variability for the recorded manual and computerized AIF and VOF signal selections. The AIF arrival time is important for correct determination of t_{max}, whereas the area under the AIF/VOF curves has a direct impact on computed CBV/CBF values. The results indicate good agreement in the AIF arrival time, but the areas under the AIF/VOF curves show higher variability.

Conclusion: We have designed an automatic AIF/VOF selection algorithm and implemented it into clinical environment, where it is a part of perfusion scans and is expected to reliably deliver estimates of AIF/VOF needed for quantitative perfusion maps. This study indicates high robustness of the implemented algorithm. Generally, only 1-2 cases out of 30 are unacceptable. The most typical confounding factors are patient motion and noise, followed by the bulk-flow artifacts in regions representing large vessels in the EPI data. Spatial shifting of the signals in those regions due to off-resonance effects during bolus peak is also often a problem. From this study we conclude that a robust implementation of an algorithm for automatic selection of the AIF/VOF signals can perform comparably well as a human operator, but 100% reliability is difficult to achieve - due to unpredictable artifacts in the acquired data. Therefore, while using this algorithm in daily clinical routine, we still monitor every case to ensure that patient care is not compromised.

Acknowledgements: This work was supported in part by the NIH (2R01EB002711, 1R01EB008706, 1R21EB006860), the Center of Advanced MR Technology at Stanford (P41RR09784), the Lucas foundation and the Oak foundation.

References: **[1]** M. Mlynash et al.: AJNR Am J Neuroradiol. 2005 Jun-Jul;26(6):1479-86; **[2]** Albers, GW et al.: Ann Neurol. 2006 Nov;60(5):508-17