Ensemble Classifier for Combining Stereo Matching Algorithms

Aristotle Spyropoulos, Philippos Mordohai

STEVENS INSTITUTE of TECHNOLOGY
THE INNOVATION UNIVERSITY®

PROBLEM STATEMENT

- Stereo algorithms have different strengths and weaknesses
- Instead of trying to further optimize one approach, we propose to combine multiple matching algorithms via an ensemble classifier trained on stereo pairs with ground truth disparity maps
- Agreement and disagreement between matching algorithms is explicitly modeled, unlike conventional approaches that treat them as independent

STEREO MATCHERS

Basic Matchers

- SAD, SSD, Sobel, ZNCC, SNCC, Census, Shiftable Windows

Advanced Matchers

- MRF, rSGM, Fast Cost-Volume Filtering, ELAS, DAISY, Superpixels

All matchers are treated as black boxes that output only a disparity map

MATCHER SELECTION

Matchers are sorted by descending number of correct disparities. The top matcher is selected and its correct disparities excluded. The process is repeated on the remaining matchers and pixels until eight *active* matchers have been selected.

Iteration:	1		2		3		4		5		6
Method	Correct Pixels	Method	Correct Pixels	Method	Correct Pixels	Method	Correct Pixels	Method	Correct Pixels	Method	Correct Pixels
SUPER-rSGM5	11,558,271	MRF-Cost	416,122	RHEM	147,758	DAISY	90,369	SH-ZNCC21	41,837	SH-SOB21	23,989
rSGM5	11,344,812	RHEM	385,280	DAISY	131,220	SH-ZNCC21	69,380	SH-SNCC5-21	41,810	SH-SAD7	21,856
MRF-Cost	11,203,915	SH-SNCC5-13	347,021	SH-ZNCC21	119,223	SH-SNCC5-21	62,040	SH-SOB21	36,025	SH-SAD9	21,670
DAISY	11,181,188	SH-SNCC5-21	342,947	SH-SNCC5-21	115,227	SH-ZNCC13	57,488	SH-ZNCC13	34,983	SH-SAD11	21,483
rSGM2	10,967,150	SH-SNCC5-11	341,270	SH-ZNCC13	99,089	SH-SOB21	52,431	SH-CSAD5-13	34,684	SH-SAD13	21,439
SNCC5-21	10,713,341	SH-ZNCC21	336,724	SH-SOB21	94,074	SH-ZNCC11	52,427	SH-CSAD3-13	34,345	SH-CSAD3-9	21,342
SNCC5-19	10,708,091	SH-ZNCC13	332,715	SH-SNCC5-13	93,986	SH-SNCC5-13	51,461	SH-SNCC5-13	34,186	SH-SOB13	21,323
SNCC5-23	10,682,715	DAISY	325,689	SH-CSAD5-13	91,902	ZNCC21	50,366	SH-SNCC3-13	33,830	SH-SSD5	21,278
SH-SNCC5-13	10,628,386	SH-SNCC3-13	324,088	SH-ZNCC11	90,199	SH-CSAD5-13	50,158	SH-CSAD5-11	32,373	SH-CSAD3-13	21,236
SH-SNCC5-11	10,614,895	SH-ZNCC11	322,875	SH-SAD13	89,985	SH-CSAD3-13	49,125	SH-ZNCC11	32,321	SH-SSD7	21,156
SNCC5-15	10,589,599	SH-CSAD5-13	321,797	SH-SNCC3-13	89,091	SH-SNCC3-13	48,368	SH-CSAD3-11	32,236	ELAS	21,111
SUPER-SNCC5-21	10,553,119	SNCC5-23	320,210	SH-CSAD3-13	88,960	ZNCC19	48,128	SH-SNCC5-11	31,546	SH-CSAD3-7	21,059
CSAD5-19	10,515,660	SNCC5-21	316,002	SH-SNCC5-11	86,550	SH-SNCC5-11	48,116	SH-SNCC3-11	31,506	SH-SAD5	20,894

122 matchers were considered with a combined coverage of 99.55%. The top matcher has a coverage of 91.9%, while the top eight matchers cover 98.57%.

Matcher 1 Matcher 2 Matcher 3 Matcher 4

EXPERIMENTAL VALIDATION

KITTI Vision Benchmark Suite

Ensemble	a_i	DD	a_i DD	LRC	a_i LRC	TS	# of Features	Out-Noc	Out-All
A	✶						5	6.81%	8.69%
B	✶	✶					11	6.54%	8.45%
C		✶					7	6.82%	8.70%
D	✶	✶				✶	17	6.94%	8.70%
E	✶		✶			✶	11	6.49%	8.38%
F	✶	✶	✶				17	6.65%	8.52%
G	✶			✶			11	6.73%	8.64%
H	✶				✶		11	6.71%	8.63%
J	✶			✶	✶		17	6.58%	8.50%
K	✶			✶	✶		18	6.68%	8.59%
L	✶		✶	✶	✶	✶	23	6.42%	8.33%
M		✶	✶	✶	✶		24	6.69%	8.56%
N	✶	✶	✶	✶	✶	✶	30	**6.42%**	**8.32%**

Validation set error rates for various ensemble classifiers using six matchers and thirteen feature combinations

Training Set

- Stereo pairs: 97
- Pixels: ~45,000,000
- Trees: 50

Validation Set

- Stereo pairs: 97

Test Set

- Stereo pairs: 195

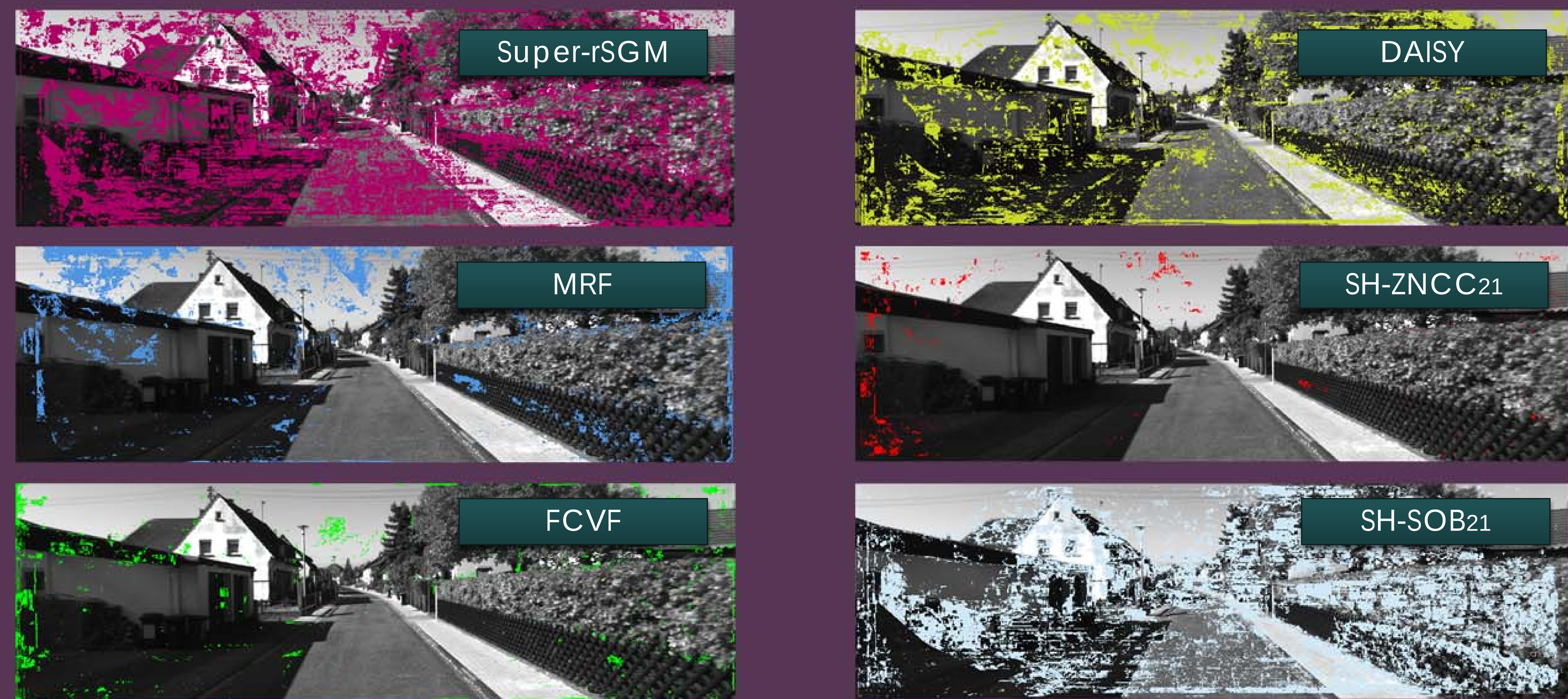
POST-PROCESSING

Using the calibrated prediction score of the winning classifier as a measure of confidence, we rejected disparities falling below a certain threshold and replaced them similarly to the paper by Bleyer at al [BMVC 2011]. Finally a 3 x 13 median filter was applied to the disparity map.

Algorighm	Matchers	Calibrated	Out-Noc	Out-All
SUPER-rSGM$_5$	8	-	8.06%	10.17%
Median	8	-	8.63%	10.64%
Majority Voting	8	-	10.24%	12.14%
N	6	No	6.42%	8.32%
N8	8	No	6.21%	8.21%
N6-C	6	Yes	6.15%	8.02%
N8-C	8	Yes	**5.82%**	7.68%
N6-CP	6	Yes	5.36%	6.87%
N8-CP	8	Yes	**5.03%**	6.48%
N8-CP (Test Set)	8	Yes	**5.34%**	6.91%

Comparison of the best matcher, SUPER-rSGM5 and median and majority voting with ensembles of six (6) and eight (8) matchers.

C = calibration, P = post-processing



Each figure highlights the pixels that selected the corresponding disparity of each matcher.

ONE-AGAINST-ALL CLASSIFIERS

Primary Matcher: The matcher whose proposed disparity is tested for correctness

Secondary Matchers: The remaining (seven) matchers whose features are used to support the accuracy of the Primary Matcher

FEATURES

AGREEMENT (a_i)

For each secondary matcher ai is equal to 1, if the primary matcher agrees in disparity, and -1 otherwise

DISTANCE FROM DISCONTINUITY (DD)

Measured as the horizontal distance from each pixel to the nearest disparity discontinuity

LEFT-RIGHT CONSISTENCY (LRC)

Defined as the absolute difference of the disparities of the left and corresponding right pixel

TOTAL SUPPORT (TS)

Number of matchers that agree with the primary matcher

FEATURE VECTOR

The full feature vector for a pixel for *m* matchers is then:

[$\{DD_{pri}\ \ LCR_{pri}\}$
$\{a_2\ \ DD_2\ \ LRC_2\ \ a_2{*}DD_2\ \ a_2{*}LRC_2\}$
$\{a_3\ \ DD_3\ \ LRC_3\ \ a_3{*}DD_1\ \ a_3{*}LRC_1\}$
...
$\{a_m\ \ DD_m\ \ LRC_m\ \ a_m{*}DD_m\ \ a_m{*}LRC_m\}$
$\{TS\}$]

The feature vector is used as input to a Random Forest classifier which calculates the likelihood of its primary matcher's disparity being correct. *m* such classifiers assign *m* scores to the proposed disparities of their primary matchers.

Once training has been completed, we perform *Classifier Calibration* using the pair-adjacent violators (PAV) algorithm. The classifier, operating on individual pixels, assigns the final disparity value of each pixel to that of the matcher with the highest posterior probability.

CONCLUSION

Despite the use of simple features, our approach is always able to surpass the accuracy of the best matcher in the active set.

ACKNOWLEDGEMENTS: The research has been supported in part by NSF award #1217797 and #1527294