IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727
PP 12-19
www.iosrjournals.org

An Improved Waytocomputedepthmapsfor Multi-View videos

[1]M.Vijayasarathy, [2] G.Balakrishnan, [3]Hariharan Ram Shankar, [4]Shreyas Prakash,
[1]*Temporary Faculty, Department of ICE, National Institute of Technology*
[2]*Director, Indra Ganesan College of Engineering*
[3]*Student, Department of ICE, National Institute of Technology*
[4]*Student, Department of ICE, National Institute of Technology*

Abstract: This paper deal swithdepthmapsextractionfrommulti-viewvi-deo.Contraryto standard stereo matching-basedapproaches,de-pthmapsarecomputedhereusingopticalflowestimationsbe tweenconsecutiveviews.WecompareourapproachwiththeoneproposedintheDepthEstimationReferenceSoftware(DERS)fornormalizationpurposesintheISO-MPEG3DVgroup.Ex-perimentsconductedonsequencesprovidedtothenormalizationcommunityshowthatthepresentedmethodprovideshighqualitydepth maps in terms of depth fidelity and virtual views synthesis.Moreover,beingimplementedontheGPU,itisfarfasterthantheDERS.

Index Terms: Depthmaps, disparitymaps, Multi-view videos, 3DVideos, 3DTV, Normalization, GPGPU, and Opticalflow

I. INTRODUCTION

Inthefuture3DTVbroadcastingscheme,manydisplaytypeswillbeavailable,leadingtomanyrequiredinput3DTVcontent.However,atacquisitiontime,itisinconceivabletocapturespecificcontentforallpossibletargeteddisplays.Thatiswhyaninter-mediatethree dimensionalrepresentationoftheviewedsceneisnecessary.Thefundamental3Dinformationattachedtomulti-viewvideosremainsthedepthmap,leadingtomulti-viewplusdepth (MVD)videos.Thesedepth mapscanbe furtherprocessedandtransformedintoother3Drepresentation,forinstanceforcod-ingefficiencypurposes.Depthmapsextractionfrommulti-viewcontentisthusamandatoryandchallengingstep,sincethequalityofthecomputedmapsimpactsalltheremainingbroadcastingchain,fromcodingtorenderingquality.Inthispaper,wepresentadifferentpointofviewofdepthmapsgenerationfromthestandardstereomatching-basedapproach,us inghighqualityopticalflowestimationalgorithms.Thecontextofnormalizationworkstowhichourmethodcanbecomparedtoispresentedin Section 2.The optical flow-based approach to depthmapsestimationisbrieflyreviewedinSection3,andresultsarepresentedinSection4.

II. MULTI-VIEWNORMALIZATIONCONTEXT

Ongoingworksforfuture3DTVcompleteframeworknormaliza-tiondealwithseveralaspects:acquisitionofmulti-viewcontent,3Drepresentation,intermediateviewsgeneration,coding/com-pressionandtransmission,etc.Themain3Drepresentationinuseis here the depth map. As such, extraction of this information has been addressed by the MPEG community under the formofaDepthEstimationReferenceSoftware(DERS).Thissoftwaretransformsmulti-viewvideosintomulti viewplusdepthvideos(MVD).Evaluationofsuchgenerateddepthmapsisperformedthroughvirtualviews synthesisusingreferencesoftware:VSRS,theViewSynthesisReferenceSoftware.
Weassumeherethesameacquisitioncontextthantheoneproposedbythe MPEGnormalizationgroup.Weconsiderthatinputvideosarerecordedbyanalignedcamerabank,withnonconvergingcameras.Theirimageplanesarethusalignedtothe samevirtual3Dplane.Wenoticethatsuchrecordingprocessisverydifficulttosetup,andthusinputimagesaregenerallycorrectedtobeperfectlyalignednotonlygeometricallyspeaking,butalsochromatically.
Depthestimation. DepthmapsestimationwithintheDERSismainlyinspiredbytheworksbelongingtothestereomatchingcommunityofthepastfewyears.Tosimplify, disparitiesareesti-matedinthreedifferentsteps [1]:

1. Localsearchofpixelmatchesalongimagelines
2. Globaloptimizationoverthewholeimage
3. Post-processing

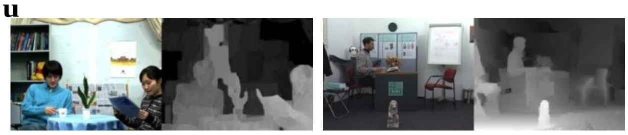
TheDERShasbeenadaptedheretofittherequirementsofthemulti-viewcontext.Insteadofusingonlytwoviews(leftandright),threeinputview(left,centralandright)areconsidered.Thedisparitymapiscomputedforthecentral view usingmotionesti-mationfrombothoftheleftandrightviews,totackleefficientlywithoccludedzonesforinstance.Implicitly,thisframework imposesthatmotionsfromleftorrightviewsbeequivalentandthusthatthethreecamerasareperfectlyaligned,thecentral onebeingatanequaldistance from the two others.Thisdepth estimation frame-workfortheDERSisillustratedonFigure1(here,forillustrationpurposes,weshowdepthmapsforthesequence*Cafe*computedwithourmethod,describedinSection3).Forinstance,thedispar-itymapforview3iscomputedusingpixelsmotionbetweenview3andviews2and4.
Figure2illustratesdepthmapsextractionresultsfortwotestse-quences:*Newspaper*and*BookArrival*.Disparitymapsareen-codedingreyscaleimages:darkvaluesindicatepixelsfarawayfromthe cameras while bright values depict near objects.Depthsmapsfor*Newspaper*arecomputedinanautomaticwaywhilemanualdisparitydatahavebeenintegratedintheestimationfor*BookArrival*.Somedisparitiesarebadlyestimated(*Newspaper*:pixelsonthebackgroundabovetheleft-mostcharacterarenotedmuchnearerthantheyshoulddo).Someinconsistenciesbetweenthemapscanalsobenoticed(*Newspaper*:onthetop-rightpartofthebackground;*BookArrival*:ontheground,tothebottom-rightside).
Thisdisparityestimationphaseiscrucial, sinceitimpactsonalltheremainingstepsofthechain.Toourpointofview,the DER

I_1andI_2.Thebrightnessofapixel$\boldsymbol{x}$inI_1shouldbeequaltothebrightnessofthematchingpixeldisplacedbyamotionvector$\boldsymbol{u}(\boldsymbol{x})$ inI2:$I_1(\boldsymbol{x})= I_2(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))$ (1)
Bylinearizingthisbrightnessconsistencyconstraintwitha Taylorexpansion, and droppingthenegligiblesecond-andhigher-orderterms, onegetstheOFC:

$$\boldsymbol{u}(\boldsymbol{x})^T \nabla I_2(\boldsymbol{x}) + I_2(\boldsymbol{x}) - I_1(\boldsymbol{x}) = 0 \qquad (2)$$

Disp1Disp2Disp3Disp4Disp5
Horn&Schunkshowedthatsolvingfor$\boldsymbol{u}$canbeperformedinan
Figure1: DepthestimationframeworkfortheDERSenergyminimizationframework [4]:

$$\bar{\boldsymbol{u}} = \underset{\mathbf{u}}{\operatorname{argmin}} E(\boldsymbol{u}) = \underset{\mathbf{u}}{\operatorname{argmin}} \left(E_{data}(I_1, I_2) + E_{prior}(\boldsymbol{u})\right) \quad (3)$$



(a) *Newspaper* (b)*BookArrival*

Figure2: ExampleofdisparitymapsextractionfortwoMPEGtestsequenceswithourmethod
Startingfromthismodel,onecansetuptheenergyformalizationasadisparitypreservingandspatiallycontinuousformu

ationoftheopticalflowproblem,basedonaL^1datatermandanisotropicTotal-Variationregularizationterm[5,6]:
Z
Didnotbehavewellenoughasis. WethuspresentinSection3softwaresimilartothespiritofDERS.ItisbasedonWerl-
$E\mathrm{data}(I1,I2)=\lambda|I2(\boldsymbol{x}+\boldsymbol{u}(\boldsymbol{x}))-I1(\boldsymbol{x})|dxdy$ (4)ΩZ
Berger’sopticalflowalgorithm[2].

ViewSynthesis. Evaluationofdisparityordepthmapscanbeper-$E\mathrm{prior}(\boldsymbol{u})=\Omega|\nabla \boldsymbol{u}_x|+|\nabla \boldsymbol{u}_y|dxdy$ (5)

formedusingthefollowingprotocol.Twononneighboringviewsandtheirassociatedmapsareconsidered.TheVSRSisusedtogenerateanintermediateview,correspondingtooneoftheac-quiredviews.Thesetwoviews—thegeneratedoneandtheac-Hereλbalancesdata and prior terms, Ωrepresentsthe imagedo-main,and$\nabla \boldsymbol{u}$isthespatialgradientofthemotionfield.Bylin-Earizingthedataterm, onegetsaconvexoptimizationproblem: Z\ Quiredone–arethencomparedusinganobjectivemetric. ThismetriccanbeforinstancethePSNR, orthemorerecentPSPNREdata $(I1, I2) = \lambda \Omega |\rho(\boldsymbol{u}(\boldsymbol{x}))| dxdy,$ (6) measure,whichaimstoconsiderperceptualdifferencesbetweentheimages.

TheVSRSusesasinputdatatwovideosandtheirassociateddis-paritymaps(whichareconvertedinternallytodepthmaps),cor-respondingtotwodifferentacquiredviews.Italsoneedsthecam-eraparametersofthesetwoviews(bothintrinsicandextrinsic).Acentralviewisgenerated,usingthecamera parameterswhich aredesired forthisvirtualpointofview andtheotherinputdata.No3DinformationisgeneratedwiththeVSRS,only2Dimages.

III. Opticalflow-Basedapproach

PrincipleandWerlberger'smethod. Basedontheobservationthatadisparityfieldestimationisnothingelsebutadensemotionestimationbetweentwoimages,weexploredanewpathtowardsrecentopticalflowestimationmethods.Theseopticalflowalgorithmsareapartofthe2Dmotionestimationalgorithms.Theirparticularitycomesfromthefactthattheyseektofindthepro-with$\rho(\boldsymbol{u}(\boldsymbol{x}))$beingtheOpticalFlowConstraintfromEquation2.Suchconvex formulationensurestheminimizertofindtheglobalminimumoftheenergyfunctional.Finally,tomaketheiralgo-rithmevenmorerobust,[2]introduceananisotropic(*i.e.*image-driven)regularizationbasedontherobustHubernorm[7]. Another extensionoftheir approachconsistsinintegration intheflowestimationnotonlythecurrentimageand the following, butalsothepreviousimage.Thegoal, intheoriginalpublication, istocopewithsingledegradedimageswithinavideo, forinstancewithhistoricalvideomaterial.

Usingoptical owinaMVDcontext. Wepresentthereasoftwarebasedonopticalflowestimationexplainedinthepreviouspara-graph,designedtodirectlyconvertmulti-viewvideostomulti-viewvideosplusdepth,called*mv2mvd*.ContrarytotheDERS,itcomputesthedisparityand/ordepthmapsforallviewsinonesinglepass,insteadofoneviewaftertheother.It'scoreusesjectedrelativemotionofscenepointswithregardstothecameras,theCUDA-based librarydevelopedinparallelwith[2].Thecore whichshouldbetheclosestpossibletothe"true"projectedmo-tion.Withsuchdefinition,theycanbeopposedtomotionesti-mationmethodsusedinthevideocodingcommunity,whichaimatfindingmotionvectorsthatminimizeanmotion-compensatedimagedifferenceinalocalsearchwindow.Forinstance,inlargetexturelessregions,ifallmotionvectorsareequivalentintermsofimagedifference,theywillfavornullvectorssincetheyareeasiertoencode,whichwillnotrepresentthetrue 2Dmotion.

WenowderivetheOpticalFlowConstraint(OFC) inamorefor-mal way,goingtoWerlberger'sformulation.Theessentialprinci-plesareexplained.Muchmoredetailscanbefoundin[3].TheOFCcomes fromthe brightnessconsistencybetween twoimagesframeworkof*mv2mvd*isdepictedonFigure3. Ononehand,dis-paritiesarecomputedfromlefttorightviews(Figure3,secondrow).Ontheotherhand,theyareestimatedfromrighttoleft (Fig-ure 3,third row).The interestof suchtwo-sidecomputation istobeabletolocateocclusionzoneswherethemotionfieldwouldbeincorrect(a pixel inone view would notbevisibleintheotherview).Infact,crosscheckisperformed[1]todetectoutlierpix-elsineachcomputeddisparitymap,whicharefinallycombined(Figure3,fourthrow)bytakingtheminimalvalueofeachdis-paritypair,toavoidtheforegroundfatteningeffectexhibitedbywindow-basedalgorithms[1].

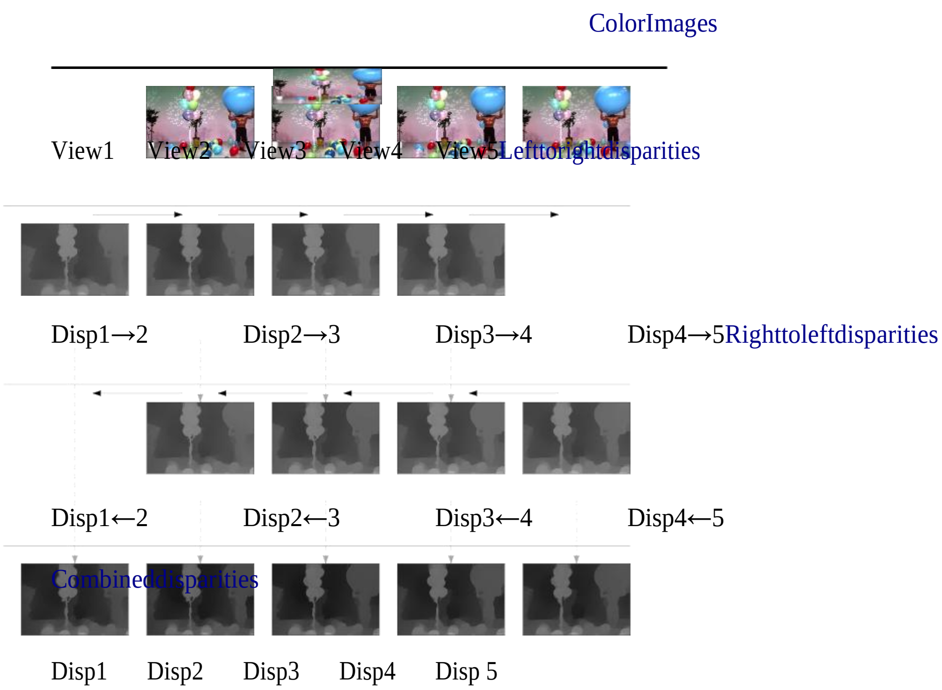


Figure3:Globalframeworkofdisparityestimationwith*mv2mvd*
(sequence*Balloons*)

IV. RESULTS

We presentin Figure4 disparitymaps extracted with ourmethodforthesequences *Newspaper*and*BookArrival*,togetherwithoriginal imagesanddisparity mapscomputedwith theDERSandprovidedtotheMPEGcommunity. Eachlineinthefigureisre-latedtooneoftheoriginalviews.Theoriginalimagesareinthefirstcolumn,disparitymapscomputedwith*mv2mvd*areinthesecondone,whileDERS-basedmapsareinthethirdcolumn.

Wecannoticethatglobally,estimateddisparitiesseemperceptu-allymorerelevantwithregardstothescenewithourapproach.Forinstance,forthesequence*Newspaper*,thebackgroundofthesceneiscorrect.WiththeDERS,numerouszoneswithdifferentdepthsappear,whilethedepthofthebackgroundisphysicallythesamewithregardstoacquisitioncameras.Asforthesequence*BookArrival*,wecannoticeagreaterspatialstabilityintheesti-mateddepths,whichappearmorenoisyintheDERScase.Thiscomesforthelatterfromtheapplicationofplanefittingonmean-shiftsegments,whichbreakthelocalspatialdepthconsistencyappliedtoeachsegment.

OnFigure5,weshowvisualdifferencesbetweenouropticalflow-baseddisparities,anddisparitiesdeducedfromadepthcamera(z-cam) acquisition of the scene.These results arepresented for thecentralofthefiveinputviewsofthe*Cafe*sequence.Howsuchz-camacquisitionhasbeenperformedisdescribedin[8].Keep-inginmindthatthesez-cam-baseddisparitiesarenotrawandhavebeeninterpolatedtofitthefullvideoresolution,itisworthnoticethatourmethodcompetesverywellwiththedepthsen-sors.Forinstance,depthcontoursseemsharperwithourmethod(allsub-images).Weareevenabletoretrievesmalldepthde-tailswithmuchlessnoise(bottomrightsub-image,forthechairpart).However,foruniformzoneswithdepthsgradients, dispari-tiesarebetterestimatedwiththez-cam(seethefurthesttableforinstance),whereourmethoddiscretizestooheavilytheresultingsignal,whilehoweverbetterretrievingdepthcontours.Noticethat

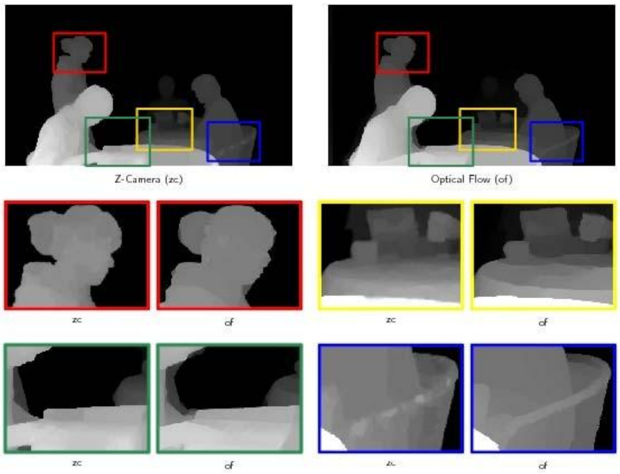


Figure5:ComparisonbetweenZ-Camera-andOpticalFlow-baseddisparities

image gammahas been modified on sub-images for visualizationpurposes. OnFigure6,wepresentevaluationresultsofourdisparitymapsintermsofvirtualviewssynthesisquality.Theevaluation protocolusedistheoneusedbytheMPEGcommunity.
Disparity maps are computedfor views*N*−1and *N*+1.These mapsareusedasinputtotheVSRS inordertosynthesizethe view*N*.Thisvirtualviewisthencomparedtotheoriginalview*N*intermsofspatialPSPNR[9]andtemporalPSPNR,withthe*Pspnr*toolprovidedtotheMPEGcommunity.Wepresentforeachsequenceandeachofthesetwomeasuresthreedifferentplots(qualitymeasureagainstvideoframenumber).TheredcurveisassociatedtodisparitymapsgeneratedbytheDERS.Theblueandblackcurvesarerespectivelyassociatedtoourmethodwith-out(MV2MVDF2)orwith(MV2MVDF3)theintegrationofthesymmetryconstraint(seeSection 3).Thedashedhorizontal linesinthefigurescorrespondtothemeanvaluesoverthewholecon-sideredsequence.
Wenoticethatfromnowon,itis difficult tobringa clearconclu-siontothisevaluationprocedure.Indeed,thequalityofsynthe-sizedviewsseemtodependmainlyontheinputdata.AnorderingofthedifferentmethodspersequenceisprovidedinTable1.ItappearshoweverthatourmethodseemtobetterbehaveinmostcasesthantheDERS.WemustalsonoticethatcomparedtotheDERS,thereisabsolutelynotemporalconsistencyenforcedinouralgorithm,sinceitseemstoprovidestableenoughresultsfromoneinstanttotheother,whichisnotthecaseofthereferencesoftware.

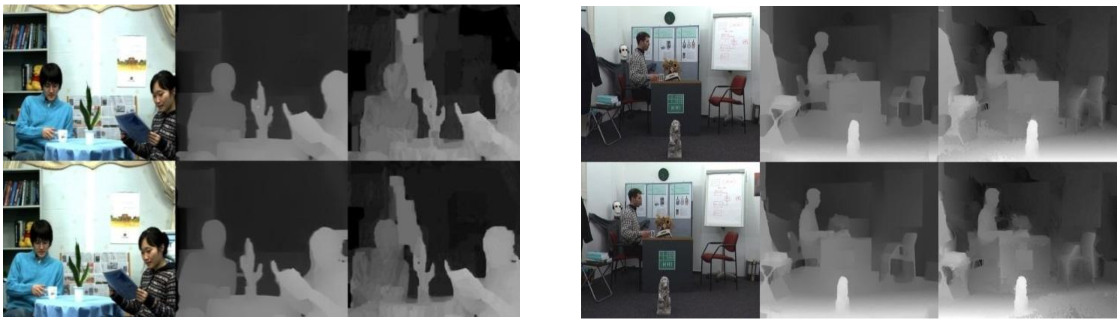
Sequence	**1st**	**2nd**	**3rd**
Newspaper	mv2mvdF3	DERS	mv2mvdF2
BookArrival	mv2mvdF2	mv2mvdF3	DERS
Lovebird1	mv2mvdF3	mv2mvdF2	DERS

Table1: Orderingofmethodsbysynthesizedviewsquality

V. CONCLUSION

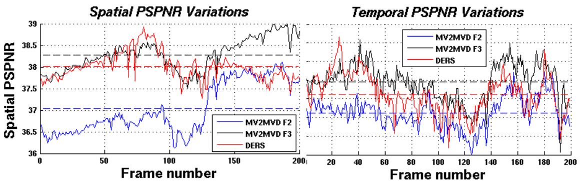
Thispaper presentsa framework for depth mapscomputation formulti-viewvideos,basedonahigh-qualityopticalflowestima-tionalgorithm.Thegeneratedmaps have been evaluated intermsofsynthesizedviewsqualityusingtheVSRSreferencesoftware.

IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661, p-ISSN: 2278-8727
PP 12-19
www.iosrjournals.org

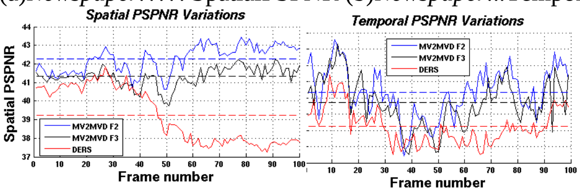


(a) Sequence*Newspaper* **(b)Sequence*BookArrival***

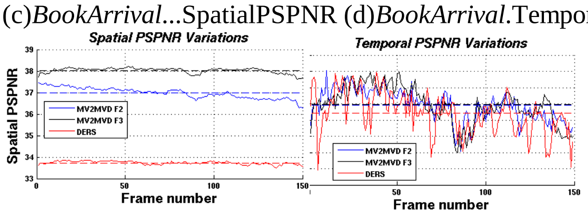
Figure4:ComparisonofextracteddisparitymapsbetweenDERS(rightcolumn)andourmethod(middlecolumn)



(a)*Newspaper*. . . . . SpatialPSPNR (b)*Newspaper*...TemporalPSPNR



(c)*BookArrival*...SpatialPSPNR (d)*BookArrival*.TemporalPSPNR



(e)*Lovebird1*. . . . .SpatialPSPNR(f)*Lovebird1*...TemporalPSPNR

Figure6: VirtualviewevaluationforNewspaper,Book *Arrival* and*Lovebird1*

Itappearsthatourmethodgivespromisingresultscomparedtomapscomputedbytheassociatedreferencesoftwarefordepthsextraction(theDERS).However,theseresultsaresubjecttomanyinterpretationsandbothmethodsarehardlycomparableforthefollowingreasons:

- Notemporalconsistencyenforcementisintegratedinoutme-thod,contrarytotheDERS.ThisisduetothefactthatsuchtemporalconsistencyintheDERScanbeinterpretedasbasedontheassumptionthatthecamerasarefixed,whichwebelieveiswaytoorestrictive.
- Wedonotproposetointegratedepthmasksduringtheestimationtoprovidemanualdataenhancingthefinalresults, contraryto the DERS.
- At writing time,we were notabletoconstrain the opticalflow estimationtobeperformedalongimagelines,asitshouldbe withcorrectlyrectifiedinputstereoormulti-viewimages.Weonlyselectthehorizontalcomponentofthecomputedmotionflow.
- Ourmethodisonlybasedontheluminancecomponentofthe inputimages,notontheRGBspace,contrary,again,totheDERS.DespitealltheselimitationswithregardstotheDERS,ourmethodisabletocomputedepthmapstotallyrelevantintermsofvirtualviewssynthesis.Moreover,beingimplementedontheGPU,itisfarfasterthattheDERS.Thecomputationaltimecanbereducedfrom15timesto150timesdependingonthemethodusedtocom-putethedisparitieswiththereferencesoftware.And lastly,onaneaseofusevision,ourmethodcomputesmapsforallviewswhenaDERSexecutionisonlyvalidforasinglevie

w, andhastoberunindependentlyforallofthem.

REFERENCES

[1]. DanielScharsteinandRichardSzeliski,"Ataxonomyandevaluationofdensetwo-framestereocorrespondencealgo rithms,"*InternationalJournalofComputerVision*,vol.47,no.1-3,pp.7–42,April2002.
[2]. ManuelWerlberger,WernerTrobin,ThomasPock,AndreasWedel,DanielCremers,andHorstBischof,"Anisotropichuber-L1opticalflow,"in*ProceedingsoftheBritishMachineVisionConference(BMVC'09)*,London,UK,2009.
[3]. WernerTrobin,*Local,semi-globalandglobaloptimizationformotionestimation*,Ph.D.thesis,GrazUniversityofTech-nology,December2009.
[4]. BertholdK.P.HornandBrianG.Schunk,"Determingopticalflow,"*ArtificialIntelligence*,vol.17,pp.185–203,1981.
[5]. NilsPapenberg,AndrésBruhn,ThomasBrox,StephanDidas,andJoachimWeickert,"Highlyaccurateopticflowcomputa-tionwiththeoreticallyjustifiedwarping,"*InternationalJour-nalofComputerVision(IJCV)*,vol.67,no.2,pp.141–158,April2006.
[6]. ChristopherZach,ThomasPock,andHorstBischof,"Adual-itybasedapproachforrealtimeTV-L1opticalflow,"in*Pro-ceedingsoftheDAGM-Symposium*,2007,pp.214–223.PeterJ.Huber,*RobustStatistics*,JohnWileyandsons,1981.
[7]. Eun-KyungLee,Yun-SukKang,Jae-IlJung,andYo-ShunHo,"3-dvideogenerationusingmulti-depthcamerasystem,"ISO-MPEGInputdocumentm17225,GwangjuInstituteofScienceandTechnology(GIST),Kyoto,Japan,2010.
[8]. E.Gershikov,E.LaviBurlak,andM.Porat,"Correlation-basedapproachtocolorimagecompression,"*SP:IC*,vol.22,no.9,pp.719–733,Oct.2007.