NEW METHODS FOR WIDE-BASELINE IMAGE INTERPOLATION

by

JEFF WOOD

Presented to the Faculty of the Graduate School of
The University of Texas at Arlington in Partial Fulfillment
of the Requirements
for the Degree of

MASTER OF SCIENCE

THE UNIVERSITY OF TEXAS AT ARLINGTON
December 2016

To my mother Sari and my uncle Luis Alfredo (my father figure)

who set the example and who made me who I am.

ACKNOWLEDGEMENTS

I would like to thank my supervising professor Dr. Vasant Prabhu for constantly motivating and encouraging me, and also for his invaluable advice during the course of my doctoral studies. I wish to thank my academic advisors Dr. Jonathan Bredow, Dr. Chien-Pai Han, Dr. Harold Sobol, Dr. Saibun Tjuatja and Dr. Stone Tseng for their interest in my research and for taking time to serve in my dissertation committee.

I would also like to extend my appreciation to Nortel Networks for providing financial support for my doctoral studies. I wish to thank Meng Yee, Robert Hunt and Dr. Michael Maragoudakis with Wireless Network Engineering for their support and encouragement. I am especially grateful to Mazin Al-Shalash for his interest in my research and for the helpful discussions and invaluable comments. I wish also to thank Dr. Yaser Ibrahim, Dr. Mini Vassudevan and Kal Mustafa for taking the time to critically evaluate this manuscript.

I am grateful to all the teachers who taught me during the years I spent in school, first in Palestine, then in Iraq and finally in the Unites States. I would like to thank Dr. Saleh Al-Araji for encouraging and inspiring me to pursue graduate studies.

Finally, I would like to express my deep gratitude to my brothers who have encouraged and inspired me and sponsored my undergraduate and graduate studies. I am extremely fortunate to be so blessed. I am also extremely grateful to my mother,

sister and wife for their sacrifice, encouragement and patience. I also thank several of my friends who have helped me throughout my career.

November 15, 2003

ABSTRACT

NEW METHODS FOR WIDE-BASELINE
IMAGE INTERPOLATION

Jeff Wood, M.S.

The University of Texas at Arlington, 2016

Supervising Professor: Farhad Kamangar

The increasing demand for voice and data capacity has been the primary motivation for cellular and PCS evolution. Cohere detection is the most power-efficient scheme that is capable of providing substantial improvement in system capacity over noncoherent and differentially coherent schemes. For this reason, reverse link coherent detection is being considered as the framework for third generation wireless communication systems. In mobile communications, however. rapid fading may preclude a good estimate of the channel phase required to achieve coherent demodulation. This may lead to serious degradation to system performance. This dissertation investigates the capacity and error-rate performance of coherent systems with imperfect carrier recovery. These systems are known as partially coherent systems.

Partially coherent systems have not received thorough investigation in the literature. Most of the previous work has been focused on the analysis of performance for BPSK over AWGN channels. Upper bounds on bit error probability have been derived, but found to be very conservative for the range of carrier phase jitter vari-

ance of practical interest. The error performance for partially coherent QPSK has not received much attention. Furthermore, the performance of partially coherent systems over multipath fading channels with diversity has not been studied.

In this dissertation, several upper and lower bounds on the error performance of partially coherent systems are derived by the application of Jensen's inequality and the isomorphism theorem from the theory of moment spaces assuming that the carrier phase error could have either Tikhonov or Gaussian distribution. An analytical method based on Gram-Charlier series expansion is also developed for the computation of the error probability and signal-to-noise ratio distribution of partially coherent systems over fading channels with diversity.

The application of partially cohere systems for CDMA mobile cellular communication is also investigated. performance impairments due to thermal noise, multipath fading, multiple access interference and self-noise are included in the analysis. A design criterion for adding weak signals with equal gain combining is established when the multipath intensity profile is nonuniform.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table Page

COMMONLY USED SYMBOLS AND NOTATION

Symbol	Description
$\mathbf{v}$	*Vectors* in *lowercase* bold
v_a	a-component of vector $\mathbf{v}$
$\mathbf{M}$	*Matrices* in *uppercase* bold
$M_{r,c}$	Entry in row r and column c of matrix $\mathbf{M}$
$\mathbf{m}_c$	*Vector* occurring in column c of matrix $\mathbf{M}$
$\mathbf{x}$	Generic 3-dimensional spatial coordinate
$\tilde{\mathbf{x}}$	Generic 3-dimensional spatial coordinate (expressed *homogeneously*)
$\mathbf{y}$	Generic 2-dimensionals image coordinate
$\tilde{\mathbf{y}}$	Generic 2-dimensional image coordinate (expressed *homogeneously*)
$\mathbf{u}$	Pixelized 2-dimensional image coordinate
$\tilde{\mathbf{u}}$	Pixelized 2-dimensional image coordinate (expressed *homogeneously*)
${}^{A}\mathbf{x}$	Generic 3-dimensional spatial coordinate in reference frame A
${}^{A}\tilde{\mathbf{x}}$	Generic 3-dimensional spatial coordinate (expressed *homogeneously*) in reference frame A
${}^{C}_{B}\tilde{\mathbf{M}}$	Change from of reference frame B to reference frame C
s	Normalizing factor applied to *homogeneous* vector $\tilde{\mathbf{v}}$ such that original $\mathbf{v} = s \cdot \tilde{\mathbf{v}}$ is recovered

${}^{D}\mathbb{S}$	Spatial reference frame D
$[\mathbf{x}]_{\times}$	Skew-symmetric matrix version of vector $\mathbf{x}$ used as *left*-operand in the *cross*-product such that $[\mathbf{x}]_{\times} \cdot \mathbf{y} = \mathbf{x} \times \mathbf{y}$
l	Epipolar line
$\mathbb{P}$	Ray (or *pencil*) of all possible vectors $\mathbf{x}$ where $\mathbf{x} = s \cdot \tilde{\mathbf{x}}$ for some value of s

CHAPTER 1

Background

Oridinarily, real-world data contains 3-dimensions. Because standard images only include 2-dimensional data, information regarding depth is lost (i.e. it is often difficult to judge distance from a single image without visual cues). *Stereovision* attempts to resolved this by finding the same point in both *stereoscopic* images (known as a *corresponding point*), and recovering the depth information. An elementry example of this occurs in stereoscopic images with relatively low distance between cameras (i.e they are righht next to each other). Objects that are *farther* away from the observer occur closer together in the stereo images, whereas objects *closer* to the camera appear appear farther appart in the stereo-images.

1.1 Change of Reference

Each view from a pair of stereo-images encompasses its own *frame of reference* (i.e. the directions of *forward* or *backward* are unique to image and may differe considerably depending on camera displacement). This requires expressing points from different frames of reference (traditionally referred to *left* and *right*) in a single reference frame. As such it is necessary to be able to express coordinates in a given reference frame in any other reference frame.

Coordinates given in ${}^{A}\mathbf{x}$ can be expressed in ${}^{B}\mathbf{x}$ by the geometric transformation:

$$^{B}\mathbf{x} = {}^{B}_{A}\mathbf{R} \cdot {}^{A}\mathbf{x} + {}^{B}_{A}\mathbf{t}$$

or

$$^{B}\tilde{\mathbf{x}} = \left[\begin{array}{c|c} {}^{B}_{A}\mathbf{R} & {}^{B}_{A}\mathbf{t} \\ \hline 0 & 1 \end{array}\right] \cdot {}^{A}\tilde{\mathbf{x}}$$

$$= {}^{B}_{A}\mathbf{M} \cdot {}^{A}\tilde{\mathbf{x}}$$

where ${}^{B}_{A}\mathbf{M}$ is also the geometric transformation necessary to transform ${}^{B}\mathbb{S}$ into ${}^{A}\mathbb{S}$.

Without calculating any new quantities, rearranging allows us to express coordinates in ${}^{B}\mathbf{x}$ in the ${}^{A}\mathbf{x}$ reference frame as:

$${}^{B}_{A}\mathbf{R}^{\intercal} \cdot ({}^{B}\mathbf{x} - {}^{B}_{A}\mathbf{t}) = {}^{A}\mathbf{x}$$

and similarly transforms ${}^{A}\mathbb{S}$ into ${}^{B}\mathbb{S}$.

1.2 Points and Lines in the Image Plane

Points in *world-space* of $\mathbb{R}^3$ are converted to points in the *image-plane* of $\mathbb{R}^2$ by *homogenization*. This occurs when a *world-coordinate* of $\mathbf{x} = [x_1, x_2, x_3]^{\intercal}$ is mapped to a *homogeneous image coordinate* of $\tilde{\mathbf{y}} = [y_1, y_2, 1]^{\intercal} = [x_1/x_3,\ x_2/x_3,\ x_3/x_3]^{\intercal}$ or a *non-homogeneous image coordinate* of $\mathbf{y} = [y_1, y_2]^{\intercal} = [x_1/x_3,\ x_2/x_3]^{\intercal}$. Points of the form $\tilde{\mathbf{y}} = [y_1, y_2, 0]^{\intercal}$ are special case of homogeneous point referred to as a *point at infinity*.

Lines in $\mathbb{R}^2$ can be represented in different contexts. The *vector offset* method calculates a line $\mathbf{s}(t)$ between points $\mathbf{y_1}$ and $\mathbf{y_2}$ as

$$\begin{aligned}\mathbf{s}(t) &= (1-t)\cdot\mathbf{y_1} + t\cdot\mathbf{y_2} \\ &= \mathbf{y_1} + t\cdot(\mathbf{y_2}-\mathbf{y_1})\end{aligned}$$

in which the line is parrallel to the vector $\mathbf{y_2}-\mathbf{y_1}$ and offset from the origin by the vector $\mathbf{y_1}$. Lines are also represented by their coefficients as $\mathbf{l} = [a, b, c]^\intercal$ where

$$\begin{aligned}\mathbf{l}^\intercal\cdot\tilde{\mathbf{y}} &= \begin{bmatrix} a & b & c \end{bmatrix}\cdot\begin{bmatrix} y_1 \\ y_2 \\ 1 \end{bmatrix} \\ &= a\cdot y_1 + b\cdot y_2 + c\cdot 1 \\ &= 0\end{aligned}$$

This definition lets us say $\tilde{\mathbf{y}}$ is located on line $\mathbf{l}$ *if and only if* $\mathbf{l}^\intercal\cdot\tilde{\mathbf{y}} = 0$. The line $\mathbf{l}$ joining two *homogeneous image coordinates* $\tilde{\mathbf{y}}_\mathbf{1}$ and $\tilde{\mathbf{y}}_\mathbf{2}$ is then calculated as the cross product of $\mathbf{l} = \tilde{\mathbf{y}}_\mathbf{1}\times\tilde{\mathbf{y}}_\mathbf{2}$.

1.3 Epipolar Geometry

Each point of of interest (also referred to as a *feature*) in a single image occurs in a 2-dimensional space at location $\tilde{\mathbf{y}} = [x, y, 1]^\intercal$. The same point in space when viewed from an image at a similar (though different) angle is referred to as a *corresponding point* with location of $\tilde{\mathbf{y}}' = [x', y', 1]^\intercal$[1]. This set of infinitley many points form a 1-dimensional subspace (also known as a *pencil*) of the 3-dimensional world space.

[1]A *change of reference* is implied between cooridinates $\tilde{\mathbf{y}} = [x', y', 1]^\intercal$ and $\tilde{\mathbf{y}}' = [x', y', 1]^\intercal$. The majority of corresponding points do not occur at the same *image coordinates* between images (i.e $\tilde{\mathbf{y}} \neq \tilde{\mathbf{y}}'$. The only way a single *world coordinate* can yield different *image coordinates*, is if a *change of reference* occurs in *world space* each time the *image coordinates* are obtained by dividing by z_{world}.

The pencil, when viewed from an image at a different angled-position, appears as a line $\mathbf{l}' = [A', B', C']^\intercal$, known as the *epipolar line*. The fact that the corresponding point (in the *angled image*) of $\tilde{\mathbf{y}}' = [x', y', 1]^\intercal$ occurs on this epipolar line is referred to as the *epipolar constraint*. It is formalized, using the previously given *line-point equality* of $\mathbf{l}'^\intercal \cdot \tilde{\mathbf{y}}' = 0$ for the *angled* image. Similarly, the corresponding point of $\tilde{\mathbf{y}}' = [x', y', 1]^\intercal$ produces an epipolar line in the *original image* of $\mathbf{l} = [A, B, C]$. The original point of $\tilde{\mathbf{y}} = [x, y, 1]^\intercal$ must lie located on this epipolar line as required by the epipolar constraint, resulting in the *line-point equality* of $\mathbf{l}^\intercal \cdot \tilde{\mathbf{y}} = 0$ for the *original image*.

When viewed in ther respective images, each point ($\tilde{\mathbf{y}}$ and $\tilde{\mathbf{y}}'$) has a pencil that coincides with that point. Since the pencils act as *directional*-vectors in 3-dimensional space, there is a unique 2-dimensional plane which contain both of these vectors, known as the *epipolar plane*. It is the intersection of the epipolar plane with the *original image*-plane and the *angled image*-plane that results in the epipolar lines of $\mathbf{l}$ and $\mathbf{l}'$, respectively. In fact, the *epipolar plane* (in each image's *coordinate systems*)[2] has the same vector form as its epipolar line. Specifically, in the *original image* reference frame $\mathbf{l} = \mathbf{P} = [A, B, C]^\intercal$, and in the *angled image* reference frame $\mathbf{l}' = \mathbf{P}' = [A', B', C']^\intercal$. This results from the fact that any *world*-point $\mathbf{x}$ lying on the *epipolar plane* $\mathbf{P}$ will result in a *homogeneous image*-point $\tilde{\mathbf{y}}$ that also lies on the plane $\mathbf{P}$. Specifically, when $\mathbf{x} = s \cdot \tilde{\mathbf{y}}$ for some non-zero value of s, then $\mathbf{P}^\intercal \cdot \mathbf{x} = 0$ implies $\mathbf{P}^\intercal \cdot \mathbf{x} = \mathbf{P}^\intercal \cdot (s \cdot \tilde{\mathbf{y}}) = 0$. Since $s \neq 0$, its true that $\mathbf{P}^\intercal \cdot \tilde{\mathbf{y}} = 0$.

In the majority of images, the sets of epipolar lines will converge at a point known as an *epipole*, denoted as $\mathbf{e}$ in the *original image* $\mathbf{e}'$ in the *angled image*.

[2]There is a single *epipolar plane* for each pair of corresponding points $\tilde{\mathbf{y}}$ and $\tilde{\mathbf{y}}'$. However, the single plane can be parameterized infinitley many ways, depending on the *frame of reference*

1.4 Fundamental Matrix

In stereo vision, points ($\tilde{\mathbf{x}}$) in one image I are related to the epipolar line (l') that contain the corresponding point ($\tilde{\mathbf{x}}'$) by the *Fundamental Matrix* ($\mathbf{F}$).

$$l' = \mathbf{F} \cdot \tilde{\mathbf{x}}$$

1.5 Intrinsic Calibration Matrix

A point $\mathbf{x}$ in the *camera-coordinate system* of $\mathbb{R}^3$ is projected to the point $\tilde{\mathbf{y}}$ in $\mathbb{R}^2$ by means of the *pinhole camera model*. The set of all $\tilde{\mathbf{y}}$ are the result of *rays* passing through the *image plane* located at $z = f$, and converging at the *optical center* as shown in the figure below:

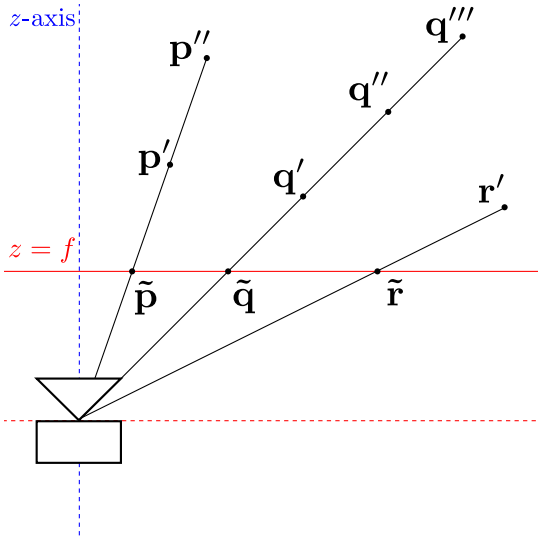

Figure 1.1. Pinhole CameraModel.

The location of $\tilde{\mathbf{y}}$ is determined by utilizing the *similarity of triangles* between $\mathbf{x}$ and $\tilde{\mathbf{y}}$. Specifically, we see that $y_1/f = x_1/x_3$ and $y_2/f = x_2/x_3$ lets us exprss the *image coordinate* $\tilde{\mathbf{y}}$ as $y_1 = f \cdot x_1/x_3$ and $y_2 = f \cdot x_2/x_3$. The point in the *image plane* of $\tilde{\mathbf{y}}$

is derived from the point $\mathbf{x}$ in *camera space* by means of the the *Camera Projection Matrix* $\mathbf{P}$ such that

$$\mathbf{P} \cdot \tilde{\mathbf{x}} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot x_1 \\ f \cdot x_2 \\ x_3 \end{bmatrix}$$
$$= x_3 \cdot \begin{bmatrix} f \cdot x_1 / x_3 \\ f \cdot x_2 / x_3 \\ 1 \end{bmatrix} = x_3 \cdot \tilde{\mathbf{y}}$$

This results in points $\tilde{\mathbf{x}}$ containing infinitley large values of x_3 being mapped to the same point $\tilde{\mathbf{y}}$ in the *image plane*. This point $\mathbf{y} = 0$ is referred to as the *principal point* in the *image plane*. This assumes the *principal point* is always located in the *image plane* at $\mathbf{y} = 0$. Projecting point $\tilde{\mathbf{x}}$ to the *image plane* with arbitrary *principal point* $\mathbf{p} = [p_x, p_y]^\intercal$ requires modifying the *projection matrix* to include *camera-specific* parameters. The *camera calibration matrix* $\mathbf{K}$ is given as

$$\mathbf{P} \cdot \tilde{\mathbf{x}} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot x_1 + p_x \cdot x_3 \\ f \cdot x_2 + p_y \cdot x_3 \\ x_3 \end{bmatrix}$$
$$= x_3 \cdot \begin{bmatrix} f \cdot x_1 / x_3 + p_x \\ f \cdot x_2 / x_3 + p_y \\ 1 \end{bmatrix} = x_3 \cdot \tilde{\mathbf{y}}$$

1.6 Essential Matrix

When coordinates from a reference frame are expressed as *normalized image coordinates* the range of possible NIC values in the corresponding image are given by the

APPENDIX A

FIRST APPENDIX NAME

APPENDIX B

JENSEN'S INEQUALITY FOR CONVEX FUNCTIONS

In this appendix, we present a procedure for improving the bounds obtained by the application of Jensen's inequality. The methiod is based on the idea of reducing the thickness of a convex region into many thinner convex regions.

B.1 Convex Functions

A real valued function f is defined to be convex over an interval $\Omega = [\alpha, \beta]$ if

$$\lambda\Phi\{x_1) + (1-\lambda)\Phi(x_2) \geq \Phi(\lambda x_1 + (1-\lambda)x_2\}. \tag{B.1}$$

If the above inequality is reversed or

$$\lambda\Phi(x_1) + (1-\lambda)\Phi(x_2) \leq \Phi(\lambda x_1 + (1-\lambda)x_2), \tag{B.2}$$

then Φ is called concave.

B.2 Jensen's Inequality for Convex Functions

Let x be a random variable with a finite mean. If $\Phi(x)$ is real-valued convex function, then

$$E[\Phi(x)] \geq \Phi\left(E[x]\right) \tag{B.3}$$

where $E[.]$ is the mathematical expectation.

APPENDIX C

UPPER BOUNDS ON MOMENTS

In this appendix, we compute upper bounds on the moments of random variables.

C.1 Computation of Bounds

Note that

$$\cos(x) \leq 1 - \frac{2}{\pi^2}x^2, \quad |x| \leq \pi \tag{C.1}$$

and

$$\cos(x) \geq 1 - x^2, \quad |x| \leq \pi. \tag{C.2}$$

REFERENCES

BIOGRAPHICAL STATEMENT

Jay V. Smith was born in Atteel, Palestine, in 1965. He received his B.S. degree from Baghdad University, Iraq, in 1988, his M.S. and Ph.D. degrees from The University of Texas at Arlington in 1993 and 1999, respectively, all in Electrical Engineering. From 1988 to 1991, he was with the department of Electrical Engineering, Baghdad University as an Instructor in the Communications and Electronics Labs. In 1995, he joined Nortel Networks, Wireless Engineering Networks as RF engineer then as a technical advisor for the design and optimization of CDMA mobile and fixed wireless access systems. His current research interest is in the area of CDMA communications for mobile cellular systems. He is a member of several IEEE societies.