Assessment of an Infrared Camera for Use as a Control Sensor for the Chemical Mechanical Planarization Process

by

Angie Shizue Nishimoto

Submitted to the Department of Electrical Engineering and Computer Science

in Partial Fulfillment of the Requirements for the Degrees of

Bachelor of Science in Electrical Engineering and Computer Science

and Master of Engineering in Electrical Engineering and Computer Science

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 21, 1999

[June 1999]

Author ....
Department of Electrical Engineering and Computer Science
May 21th, 1999

Certified by...
Duane S. Boning
Associate Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by ....
Arthur C. Smith
Chairman, Department Committee on Graduate Theses

Assessment of an Infrared Camera for Use as a Control Sensor for the Chemical Mechanical Planarization Process

by Angie S. Nishimoto

Submitted to the Department of Electrical Engineering and Computer Science on May 21, 1999, in partial fulfillment of the requirements for the degrees of Bachelor of Science in Electrical Engineering and Computer Science and Master of Engineering in Electrical Engineering and Computer Science

Abstract

Chemical mechanical polishing/planarization (CMP) is considered a black art because of the many factors which contribute to the quality of its performance. The arm which holds the wafer against the pad makes it challenging to get any readings from the wafer during polishing. The chemical and mechanical factors are difficult to separate, and even more difficult to monitor. It is this difficulty which promotes inspection into temperature sensors for the process. Both the chemical and mechanical aspects of the process are closely linked to the temperature of the process. Observation of the process temperature provides an estimate of the condition of the wafer during the polish. This is already done for average wafer end point detection (EPD). This thesis takes the idea one step further by looking into the possibility of monitoring the spatial variation in temperature, and using this as an estimator of spatial polishing conditions.

A infrared point sensor was used first to take temperature readings at specified positions during polishing. What was found was that the spatial profile of the polishing process temperature varied. This variation was dependent upon the process running, and appears to be linearly related to the post polish film thickness uniformity. It was also found that the pad break-in period could be monitored by this sensor, as the temperature signature drops off after the first batch of wafers were run. The same experiments were then conducted utilizing an infrared camera with a focal plane array of temperature sensors. This camera was used to verify the previous work. The camera also offers higher spatial and thermal resolution, increasing the accuracy of the results.

Thesis Supervisor: Duane S. Boning
Title: Associate Professor, EECS

Acknowledgments

I would like to express my sincere gratitude to my thesis advisor, Prof. Duane S. Boning. Without his constant help, encouragement and vision, this project would not have been possible.

My sincere gratitude to the NSF/SRC Engineering Research Center for Environmentally Benign Semiconductor Manufacturing for its generous support of this work.

My thanks also goes out to Taber Smith and Dennis Ouma, for being my mentors during my efforts. I am grateful for everything that you have taught me about semiconductors and not. In addition I would like to thank the other members of the Control and Statistical Metrology groups at MIT's Microsystems Technology Labs, Aaron Gower, Han Chen, David White, Eric Stuckey, Sandeep Sadashivappa, Terrence Gan, Brian Lee, Tae Park, Tamba Tugbawa, Charles Oji, and Vikas Mehrotra, for all of the knowledge which they have shared with me, and the support that they have given; Aaron for always making the computers work, Sandeep for lightening the hours, Brian for all the processing help, and everyone else for being there.

Thank you to the technical staff of the MIT Microsystems Technology Labs. Thank you for fixing the tools and for being understanding. Special thanks to Tom Tacaks and Joe Walsh for all of their insight into the CMP tool.

I would also like to thank Everette Lambert at Strassbaugh for his help with the infrared point sensor

Table of Contents

List of Figures

Chapter 1

Introduction

Chemical Mechanical Planarization is a relatively new tool in the myriad array of processes that make up a semiconductor fabrication facility. It performs the necessary task of planarizing the inter-level dielectric layer enabling higher stacks of metal wiring necessary for the complex circuits that are constructed. The process has its downfalls too, results of a small but growing knowledge base on how it actually works. To expand this knowledge there has been a great deal of work in the field to identify factors which contribute to and affect the quality of the processed wafers. This includes some work in the invention of novel sensors such as temperature sensors to overcome the processing challenges inherent in CMP.

1.1 Background

Although chemical mechanical planarization (CMP) is widely used, it is still not well understood. The coupling of mechanical and chemical abrasion at the surface of the wafer is believed to be important. In this thesis, we are particularly interested in the increased heat on the pad surface which is observed during the polishing process. Understanding what factors contribute to the observed heating of the pad may make clearer the principal mechanism of CMP. The temperature of the pad, being so strongly coupled to the process conditions and mechanisms should be a good estimator of the condition of the wafer surface during a polish. By observation of the temperature of the process, data may be acquired which would provide information about the polishing status in time and polish uniformity in space. To explore the possibilities that this temperature element can bring, an exploration into the feasibility and applicability of an infrared camera as a spatial temperature sensor for the CMP process has been conducted.

1.2 The CMP Process: Motivation

Chemical mechanical planarization is used in semiconductor fabrication to planarize topography on the wafer or remove undesired material from the surface of the wafer. This planarization is necessitated by the fact that as the constraints on line width resolution get tighter with the shrinking dimensions of devices, the depth of focus of the lithography patterning systems is becoming increasingly shallower. This means that variation in the height of the wafer surface results in substandard patterning. Without the planarization step in the process flow, the upper layers of the wafer become raised where they pass over underlying metal layers. This increases the variation in the topology of the surface targeted for patterning which then increases the non-uniformity of the distance from the patterning equipment to the wafer surface. This causes the pattern to be blurred in some areas. CMP is used in this case to planarize the inter-level dielectric (ILD) layers and enable finer line patterning as described in Figure 1.

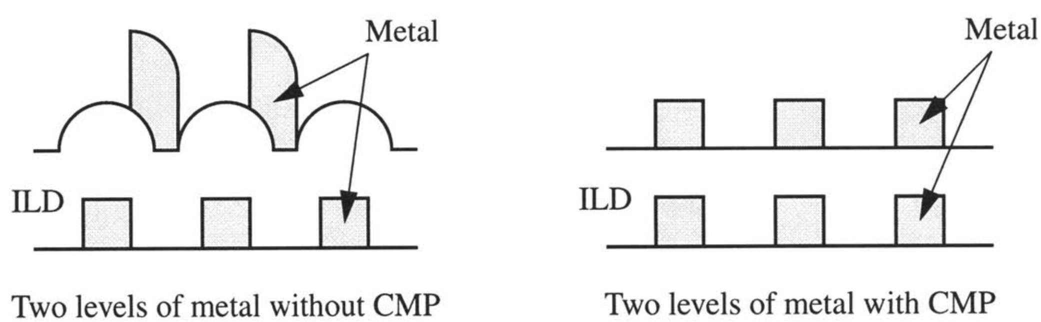

Figure 1. CMP enables higher levels of logic by planarizing ILD layers.

CMP has also become necessary as the industry moves to copper as an interconnect metal to replace aluminum. Copper as a metal has less of a tendency towards electromigration than aluminum does which means that thinner lines of copper are more reliable under the same operating conditions. Copper also has a lower resistivity meaning less loss of energy in the copper interconnects, and a lower delay along the lines. Aluminum, being an older process, has the benefit of being readily patternable by wet etch or by plasma/dry etch. The difficulty of copper compared to aluminum is that there are no good fine-feature plasma etch processes for copper patterning. The process which enables copper patterning is damascene; a process also used to produce the plugs of metal which connect the wires in one layer to those in another. In this processing scheme, lines or vias are etched into the ILD layer and an excess of copper, or other metal, is deposited onto the surface filling the lines. The overburden of copper is then polished off by CMP leaving inlaid lines or vias of copper and ideally a completely planar surface. A schematic of the damascene process is shown in Figure 2. The need to pattern copper in this way has driven the current demand for a better understanding of the CMP tool and the process itself.

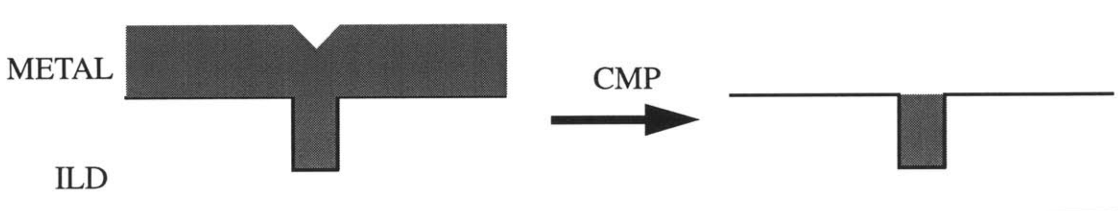

Figure 2. Copper damascene scenario.

1.3 The CMP Process: Mechanism

To accomplish efficient planarization of the wafer surface both a mechanical and a chemical component are utilized in the CMP process. Either component alone is much less efficient. The mechanical component is slow, and the chemical even slower alone, but together they work in concert. The chemical component appears to “soften” the surface of the wafer, enabling much more rapid abrasion. The abrasion of the “softened” areas exposes more material to react with the chemicals in the slurry. In the case of metal polishing, the slurry is also necessary to prevent replating of the metal back onto the surface of the wafer.[1]

A schematic of the tool is shown in Figure 3. The wafer is held in a chuck which is rotating against a larger platen head. Both the chuck and the platen are rotating in the same direction at approximately matched speeds which results in uniform and unidirectional relative velocity vector at all points on the wafer surface. The speed of the table is one of the most prominent components in the mechanical side of the process. It greatly influences both the removal rate of the process and the uniformity. The other process

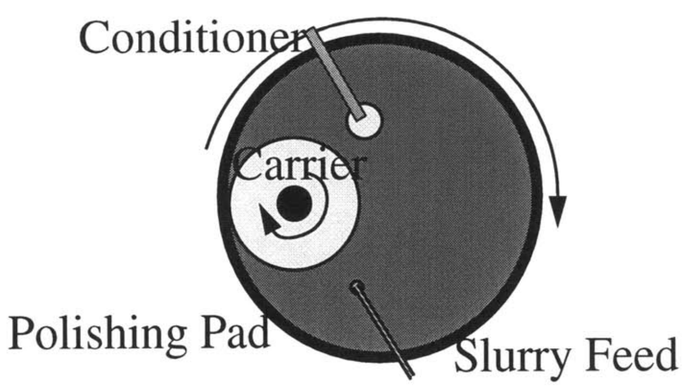

Figure 3. Top view of chemical mechanical planarization tool.

parameter which appears to be highly influential is the down force exerted by the carrier head to press the wafer against the polishing platen. This is commonly referred to as the down force. Some wafer chucks will also have a secondary mechanism for holding the wafer itself against the pad, providing back pressure from the chuck itself to the wafer.

The two main consumables utilized in the CMP process are the pad and the slurry which do the polishing. The platen of the tool is covered with the abrasive pad which is typically constructed of a polyurethane material. The choice of this pad can greatly influence the quality and speed of the polishing. Some pads are manufactured as composites of several pads stacked one upon the other to provide different mechanical bending characteristics. When it is first installed, the new pad polishes at a high rate until it becomes broken in and the removal rate settles down. As it continues to wear, it loses the ability to polish the wafer and needs to be reconditioned periodically.

During polishing, a slurry is dripped onto the pad. This slurry contains the chemicals that enable the chemical component of the polishing. It also contains particles of abrasive material. These may be made up of many things, but common abrasives are seria, silicon dioxide or alumina. The size of these particles is highly regulated, and their characteristics can influence the quantity of microscratches and defects generated by CMP. The content of the slurries is largely proprietary and matched very closely to the material being abraded on the wafer surface. The rate of polish typically follows Preston's Equation (1) which relates the removal rate to the pressure exerted on the surface by a constant K_p.[2]

$$Rate = K_p \cdot P \cdot \frac{\Delta s}{\Delta t}$$

In an ideal case where the pad is completely rigid and the surface completely flat, this would result in completely uniform removal across the wafer. The pad, however acts elastically, and the wafer is typically patterned so there is ratio of raised area to lower area which can be viewed as the density of the pattern which also redistributes the pressure of the pad. Over a large area this density averages out, but in a localized area the pressure is greater if the density of the raised area is small. A way to visualize this is to think of pillars holding up a tent. If there are a large number of large pillars, (raised surfaces on the wafer) the weight of the tent (the polishing pad) is distributed among them and each holds up only a small amount of the weight, but if there are only a few pillars or these are very small, the pressure is greater on each pillar. This non-uniform pressure distribution leads to a non-uniform polish in which the areas of low up area density polish much more quickly than the areas with a large amount of raised area. This is one of the challenges of the CMP process because it is intrinsic to the mechanics of the system.

1.4 Difficulties in Monitoring the CMP Process In-Situ

Despite its usefulness in many areas of semiconductor fabrication, the CMP process still suffers from its youth and lack of understanding. The main difficulty is the inability of the process to uniformly polish or planarize a wafer. To begin with, the wafer is not completely flat, and polishing can accentuate this curvature. Depending on the process conditions under which the wafer is polished, the polish rate and therefore the final thickness of the material can vary radially. This is referred to as wafer-level variation. It has also been observed that the final topography of the wafer can be related to the original pattern underlying the material being polished. This variation mentioned in the previous section as non-uniformity due to varying pressure distribution is referred to as die-level or pattern dependent variation. This variation is also very process dependent. Work has been done to deconstruct this variation into its component pieces. There have been several models put forth to explain the die level variation as a function of both process conditions and the pattern densities of the wafers being polished.[3]

Both wafer-level and die-level variation need to be monitored, but they can only be identified through measurements of the surface film thickness. This measurement, however is done optically or by using profilometry, and can only be done ex-situ, after the wafer has been polished and run through a variety of clean steps. In addition the clean steps often remove a quantity of the polished material complicating the measurement process. During the polishing process, however the wafer is inaccessible optically or physically, though there have been attempts to get around this limitation.

1.5 Previous Work in Field

There has been a great deal of previous work in the field. This is due to the huge surge of interest which accompanied CMP's introduction to the semiconductor community. Many techniques have been addressed as possible methods for dealing with the difficulties in the process. One field, variation monitoring, has become a large field of interest both from the point of view of the manufacturers of integrated circuits and from the perspective of the CMP tool suppliers. Current work at the Massachusetts Institute of Technology's Microsystems Technology Laboratories seeks to identify and model the variation which results from polishing. Variation on a wafer after polishing can be attributed to many sources. Two of these are wafer level non-uniformity, and run-to-run variation. This vari-

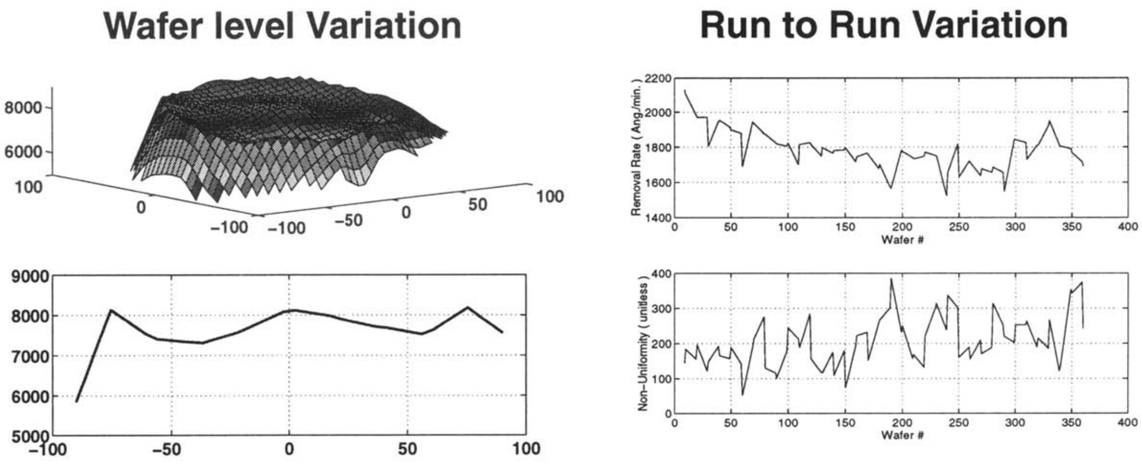

Figure 4. Types of variation in CMP processing

ation can be very significant, as shown in Figure 4. The chip makers demand better controlled processes to increase their yield by increasing the reliability and repeatability of the finished die. Both the tool suppliers and the chip manufacturers have begun looking into sensors that can be used in-situ to determine the state and progress of a polishing process. These sensors are typically used to monitor two parameters. The first parameter is removal rate. This parameter is a strong function of the process conditions and the state of the consumables. A large change in this rate would be both an indicator that the process is drifting, and also a warning that the surface film thickness is not changing at a constant rate. These are geared typically at ILD polish, where there is a specification on how much of the dielectric layer must remain after the planarization step is completed. A second parameter which is typically monitored is end point. End point is found when polishing back one material to reveal another. The moment the process hits end point is the moment when the overburden of material is removed to reveal the highest points of the underlying

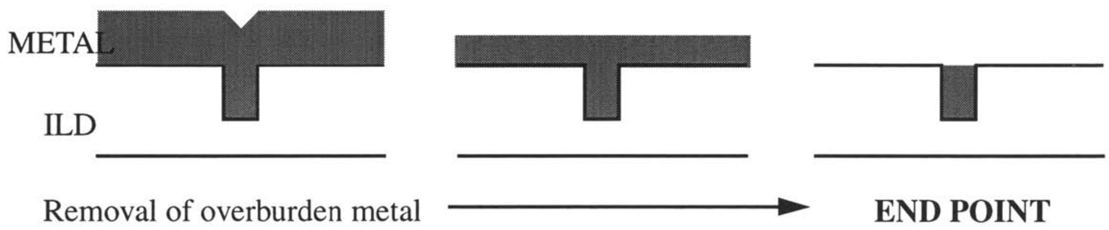

Figure 5. Ideal graphical definition of end point.

material. See Figure 5. EPD is typically used for shallow trench isolation (STI) processes, tungsten plug inserts, and for copper damascene processes. Accurate end point detection (EPD) is especially important in the metal scenarios because residual metal left on the surface of the dielectric layer can cause shorts between lines.

There have been a large variety of sensors that have come into existence to monitor the process removal rate, and to detect end point. They range from electrochemical to optical, and extend even to thermal readings used as estimators for the process conditions. For this work, we will concentrate on the efforts to develop a temperature sensor.

There has been previous work to determine that temperature is a good process parameter to monitor for EPD.[4] It was found that the breaking through of one layer to another and the subsequent change in the coefficient of friction between the pad and the wafer resulted in a drop in temperature. This drop could be monitored and used as an indication that the process had reached end point. This is one sensor that is used currently in industry. It is not very robust, however. It only reads the average endpoint as measured at whatever radius of the pad the sensor is positioned to look at. In addition, the drop in temperature is not well enough understood to determine the precise moment of endpoint, and the variation on the wafer is substantial enough that "end point" is still not used as the point when polishing stops. A significant "over-polish" step is necessary to be sure that all of the excess material has be removed at all dies on a wafer. In this scheme we are still relying on constant times, unable to correct for varying process conditions and polish rates.

Another theory is that removal rate and temperature can be correlated.[5] If this can be proven, a temperature sensor would give in-situ information that could be used to update the process conditions keeping a wafer in specifications. With an infrared camera the entire pad surface could be imaged.[6] The temperature of the pad could be monitored spatially, meaning that it would give an in-situ estimator of the polish rate uniformity across the wafer.

1.6 Infrared Temperature Profilometry

IR profilometry is the usage of spatial temperature data to infer the condition of the wafer surface with regards to the spatial uniformity of the polish. This is done by monitoring the temperature of the pad at different radii. These measurements can then be used to construct a temperature profile for the wafer which is dependent upon the film being polished and the polishing conditions. An example is shown in Figure 6 of a probable temperature profile. The concentration of temperature is greatest in the center where more

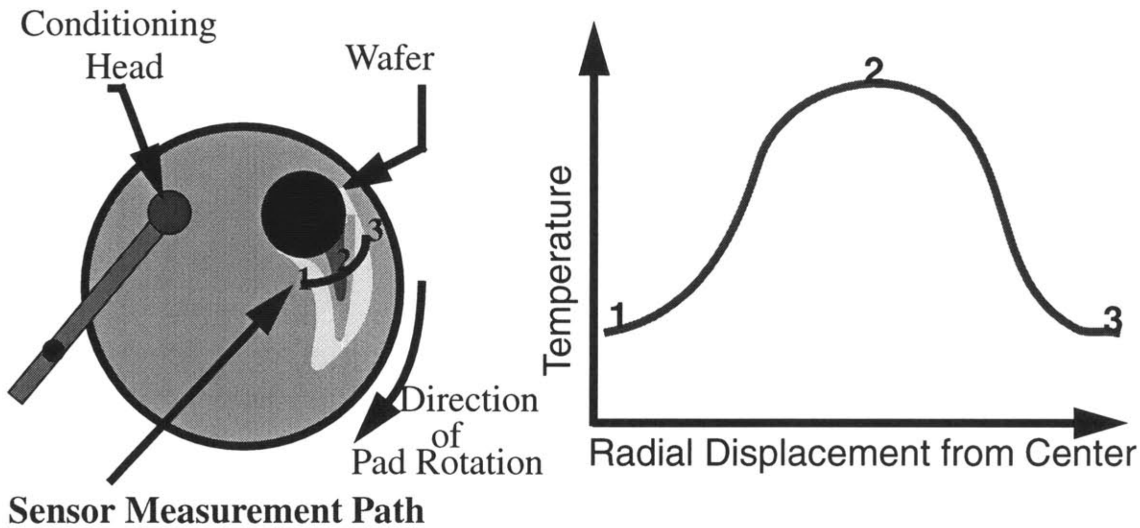

Figure 6. Idealized temperature profile of polishing wafer.

work is being done and cooler at the edges. It is this profile which can be used as an estimator for the polish uniformity of the polishing wafer.

1.7 Scope and Objectives of Thesis

This thesis will illustrate the possibilities for utilizing a temperature sensor as an in-situ process monitor for the CMP process. As a monitor and estimator of removal rate and end point, temperature offers information of the process conditions, and the progression of the run. This is true not only on an average over a large area of the polishing surface, but spatial temperature variations will be shown to be important as estimations of the quality and uniformity of the polishing process. This goes beyond the standard of the industry which utilizes only average temperature measurements to identify global characteristics of the process.

This work covers two experiments which were conducted at the Microsystems Technology Laboratories at the Massachusetts Institute of Technology. The first was an adaptation of an existing IR temperature sensor for use in monitoring the temperature of a polishing wafer at various points on the polishing pad. The second experiment was conducted to elaborate upon the spatial temperature readings by incorporating an infrared camera which would provide spatial temperature information which could then be used to evaluate the process progression.

Chapter 2

Infrared Point Sensor

The proof of concept for the idea of in-situ infrared profilometry was the implementation of an infrared point sensor as a thermal monitor of a specific pad radii during polishing. The point sensor was used to monitor the temperature at various positions on the polishing pad during polishing periods for several different process conditions. The results were promising. They show that the temperature of the pad can be related to the quality of the polish, the process settings, and to the condition of the consumables.

2.1 Experimental Setup

The first experiment utilized the single point temperature sensor which accompanied the Strassbaugh 6EC tool. The goal of this set of experiments was to determine the degree to which temperature could be monitored and correlated with the uniformity of the wafers being polished upon the tool. To correlate spatial uniformity of temperature with the spatial uniformity of the wafers, temperature time traces needed to be taken at various radii of the pad. The existing temperature sensor which accompanies the Strassbaugh 6EC tool was adapted to the task by the construction of a bracket which would mount the sensor upon the carrier head of the 6EC tool and point it directly downward towards the pad surface. Since four inch wafers were polished, five radii of the polishing pad were chosen. The first radii was chosen to be aligned with the edge of the wafer's polish track, and the others were spaced at one inch intervals to cover the entire track of the wafer during polish as described in Figure 7. The five positions were marked so that the sensor could be moved from one trial to the next with consistency. The sensor used was a Raytec sensor whose output was the averaged temperature over a one inch diameter spot in front of the camera. The sensor fed directly into the computer of the 6EC tool because the sensor was intended to allow for control of the table temperature during polishing. The signal could thus be retrieved by reading the output of the tool which was fed into a spreadsheet designed by Strassbaugh for the monitoring of the process conditions in the tool. This spreadsheet was used to store the temperature at evenly spaced intervals during the polishing. An example of a temperature trace taken during the polishing of a blanket oxide film

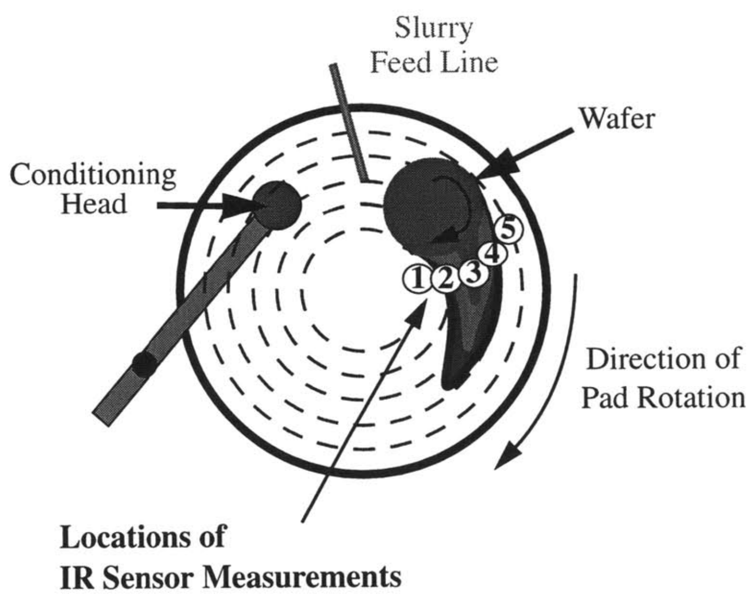

Figure 7. Top view of alignment of infrared point sensor.

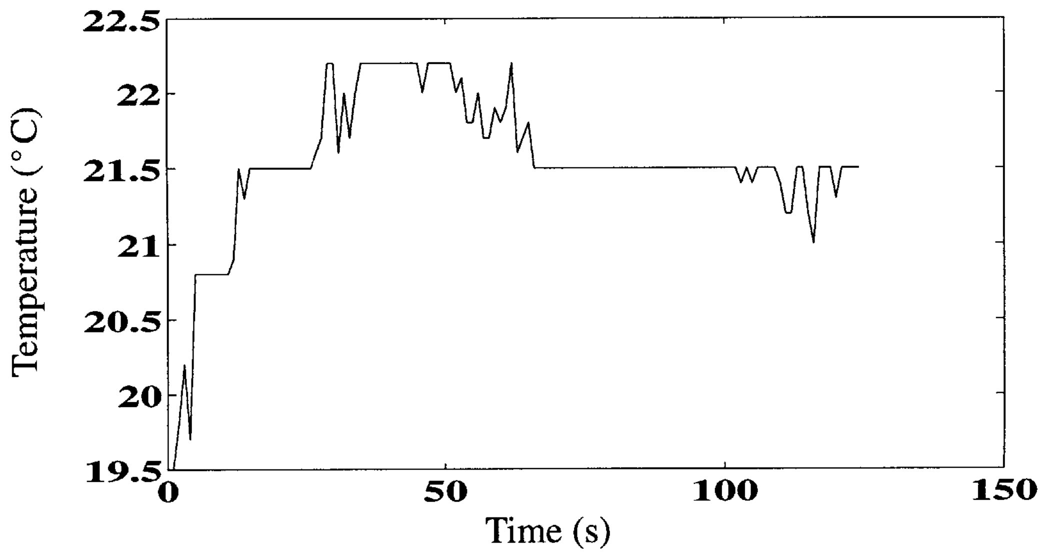

Figure 8. Point sensor temperature reading vs. time.

is shown in Figure 8. The trace begins with the onset of polishing. The temperature begins at room temperature and rises as heat is generated by the friction in the polishing. Eventually the rate of heat generation is balanced by the heat radiation from the pad as it goes around and the temperature reaches a steady state temperature. This is typical of the oxide polishing process.

With this method of temperature recording in place, a set of experiments were designed to look at the difference that process settings would have on temperature and how this variation would relate to the uniformity of the wafers which were polished.Tables 1 and 2 show the list of wafers which were run and the process conditions under which

Table 1: Process Setting Specifications

Process Recipe	Table Speed (rpm)	Slurry Flow Rate (ml/min)	Average Removal Rate (A/min)	Average Temperature
A	25	150	845	20.47
B	50	150	1037	21.89
C	25	75	750	20.57

Table 2: Wafer Specifications by Sensor Position and Recipe

Wafer Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Process Recipe	A	A	A	A	A	B	B	B	B	B	C	C	C	C	C
Sensor Position	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

they were polished. In addition to the conditions varied as shown in Table 1, all wafers were polished with a carrier speed of 25 rpm, and a down pressure of 3 psi. Wafers were polished with the sensor in each of the five positions for each of the process settings. The whole set of 15 wafers was replicated three times to reduce anomalies in the data. This resulted in a total of 45 wafers polished in random order. The randomization was introduced to prevent second order effects due to the time order in which the wafers were processed (such as drift in the tool or pad wear) form being misinterpreted as factor effects.

2.2 Results

The results of these trials were very promising. The radial temperature profiles for the baseline process (A), high speed process (B), and low slurry process (C) are shown in Figure 9.The curves shown in this graph represent the average of the profiles taken during the three replicates of the experiment. The data has also been interpolated to suggest a more

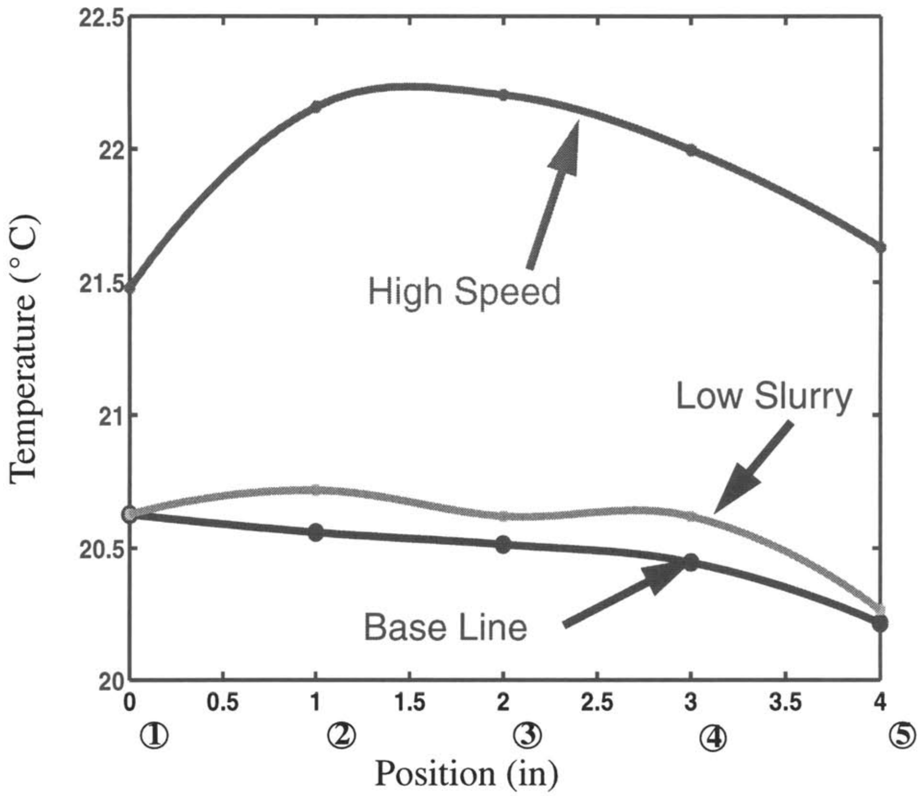

Figure 9. Interpolated average profiles for three process settings

realistic profile. Looking at these temperature signatures, the first effect which becomes evident is that each process has a characteristic shape. The high speed process has a higher average temperature and a larger curvature to its domed shape. The baseline and low slurry processes, by contrast, have lower average temperatures and smaller curvature. The latter processes are also difficult to distinguish. These observations are consistent as they confirm the significance of speed as a key CMP process variable compared to slurry flow rate which is not as significant.

Looking at the high speed process alone, we notice that there is a significant rise in temperature from the edges of the profile to the center. This supports the theory that the sensor is monitoring the aggregate work done beneath the wafer as it polishes. Because the wafer is circular, the pad contacts a larger arc in the center of the wafer than at the edges, creating more heat there. Because the low slurry process does not exhibit this shape, it is unlikely that this heat up is due to slurry transport hindrance.

When all of the data taken is taken into consideration, rather than just an average, a second key observation is available. The first repetition for all three process settings, the solid lines in Figure 10, show a significantly higher temperature than the later repetitions across all three process recipes. This is attributed to the fact that a new pad was used in this experiment. It is known that during the initial pad "break-in" period, the removal rate begins at an elevated level and drops off exponentially before stabilizing. During this period, it is necessary to polish dummy wafers to minimize the variation of the process from wafer to wafer. The sudden decrease in average temperature from the first to the second replicate set demonstrates the potential use of the radial temperature profile as a monitor for the pad break-in period. This ability to know immediately for each pad when the removal rate has stabilized could decrease the number of wafers expended when breaking in the pad.

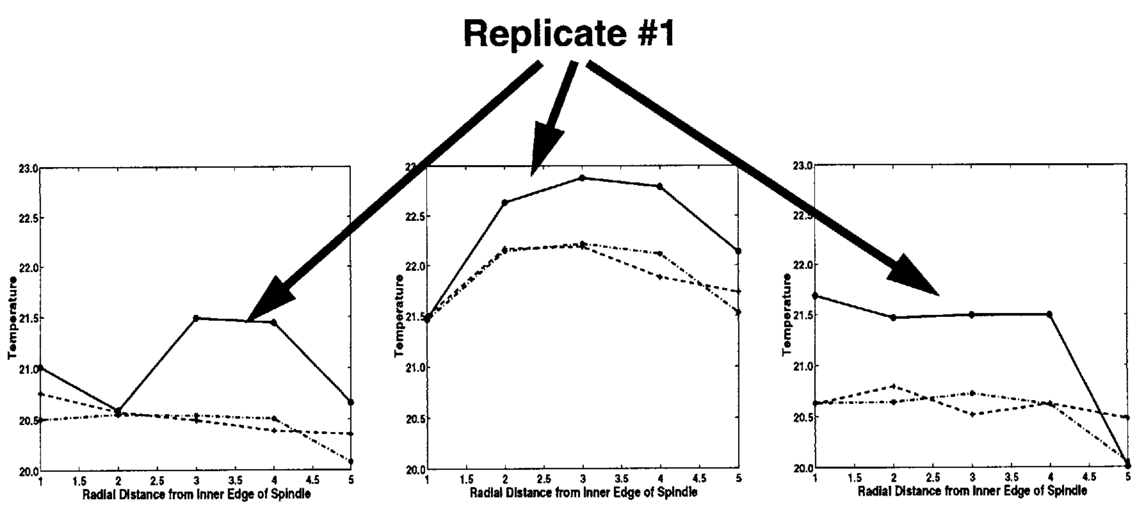

Figure 10. Profiles for each replicate of experiment taken at all process settings

The wafers were measured by taking 49 film thickness measurements ina radial pattern over the surface of the post-polish wafer. These measurements were used to calculate the non-uniformity of the polish. The metric of non-uniformity was defined to be the standard deviation of these 49 measurements divided by their mean. The non-uniformity of each temperature profile was calculated by taking the five temperatures averaged in time, one for each spatial position, and calculating their standard deviation of the five points divided by their mean.

When the variation in the temperature over the different radii of the pad was compared to the variation in wafer level uniformity of the post-polish wafer, it was also found that the variation in the temperature profile appeared to be linearly related to the variation in the final oxide thickness on the polished wafers as shown in Figure 11. The positive implication of this relationship is the possibility of modeling the non-uniformity of a polishing wafer not only as a function of process conditions, but as a function of the uniformity of an in-situ temperature measurement.

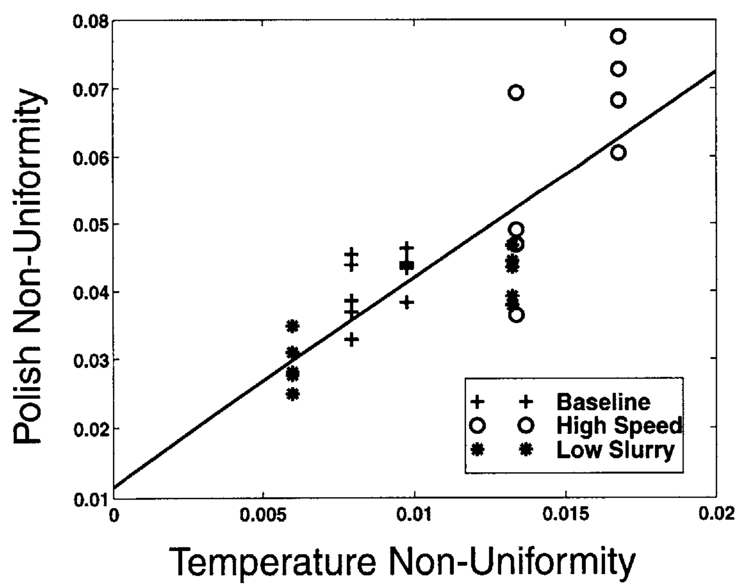

Figure 11. Comparison of the non-uniformity observed in the temperature profile and the nonuniformity evident in the post-polish wafer oxide thickness

2.3 Limitations of Point Sensor

While the point sensor was an excellent opportunity to utilize existing hardware to demonstrate the versatility of an infrared sensor, it is not an ideal implementation of a spatial infrared sensor. The point sensor only takes the average temperature over a 1 inch diameter circle. This low spatial resolution severely limits the accuracy of the sensor. The temperature reported is only an average of the temperature over the entire spot and therefore only an estimation of the temperature at the center of the spot. In addition, the point sensor could not be moved during a run, and therefore only one radii of the pad could be monitored during each wafer polish. Ideally the entire pad surface temperature would be monitored. From this data the entire profile could be obtained for each wafer, and additional information could be extracted from the heat decay across the rest of the pad surface. An infrared camera with a focal plane array of infrared sensors would give a much higher resolution image, and more accurate data about the temperature variation spatially across the pad surface. This relatively new technology would enable more accurate prediction and modeling of the relationship between temperature and wafer condition in both the case for polish uniformity and spatial EPD.

Chapter 3

Infrared Camera

With the positive results that were acquired with the point sensor, the next logical step was to eliminate some of the limitations by upgrading our equipment to an infrared camera which would give us the temperature readings for a larger area with a greater spatial resolution. The results of a second set of experiments confirm the work done with the point sensor. The new equipment also opens new doors, new possibilities for the system.

3.1 Description of Camera

The camera which was selected to be used in these experiments was an Agema/FLIR Thermovision 550. The camera is non-invasive. It does not project infrared radiation, but relies upon absorption of the natural infrared radiation being emitted from an object to measure its temperature. It has a field of view of 20° by 15° as defined by the lens of the camera. The optics focus the infrared radiation onto a focal plane array (FPA) of dimensions 320 by 240 pixels within the camera which is cooled to increase the thermal sensitivity of the array. Each pixel of this array takes an intensity reading and reports it at a regulated time interval. These intensity readings are used along with the emissivity of the body being observed and the ambient temperature to calculate the temperature of the object in view. This is all observable in real time. For measurement purposes, the accuracy of the camera is within 2% of the range. In the case of this camera the range of readings was from -20 to 250 degrees centigrade. Therefore the accuracy of the temperatures reported by the camera is within 3 degrees. The sensitivity of detection in temperature differences within an image is greater than 0.1 degree centigrade. This allows for detection of very minute temperature changes within an observation session. This high degree of sensitivity within an image is ideal for our purposes because we are most interested in the spatial variation of temperature within a run.

3.2 Design of Experiments

With the infrared camera, it was unnecessary to run the number of wafers that were required for the experiments with the point sensor because the camera was able to capture the entire profile in a single run. The goal of the experiment was to repeat the work done with the infrared sensor to compare the results previously obtained with the data from the camera. This would serve to verify the point sensor data, and also to illustrate the increased sensitivity available in the camera.

As in the previous experiment, wafers were run at three settings, one was a baseline setting, the second a high speed setting and finally a low slurry flow setting. The variations in processes are identical to those in the previous experiment; the table speed was doubled for the high speed process and the slurry flow rate was decreased by 33% for the low slurry flow process. This time only two wafers were run for each recipe. Each wafer was polished for two minutes and the entire polish time was recorded at a rate of one measurement per second by the camera.

During the experiment the camera was set up upon a tripod outside of the CMP tool looking into the side maintenance door to capture as much of the pad surface as possible.

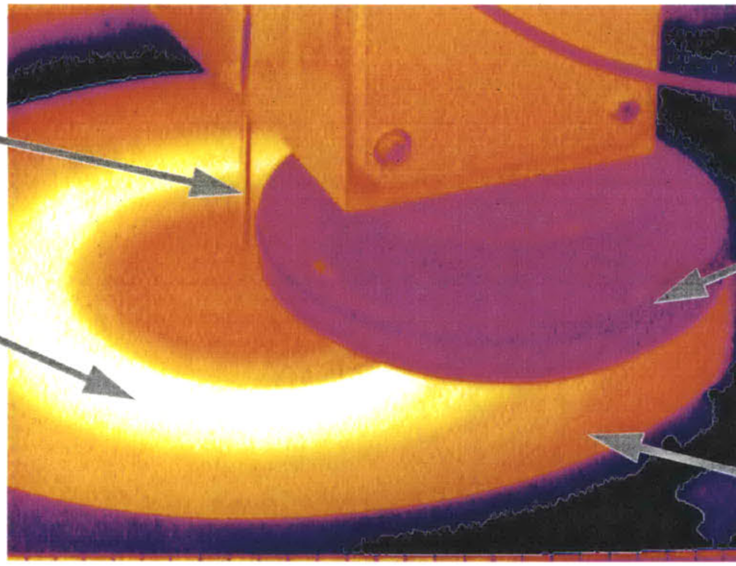

Figure 12. IR image of CMP tool polishing

An example of the image of the camera looking at the CMP tool is viewable in Figure 12. The camera stores an image like this every second. It stores not only the qualitative pictures, but also an actual intensity value for each pixel in the image. These images can then be called up later and analyzed by extraction of the temperature values. Infrared movies made up of these images were taken during each polishing run. The points extracted for analysis were taking along the lower edge of the carrier head. Their position corresponds to the positioning of the infrared point sensor in the previous experiment.

3.3 Results

Although only six wafers were run, the data files resulting from the infrared images were very large. It was a challenge to parse the data for the quantities desired. For this reason, the first observations were qualitative ones. Figure 13 is a collection if images

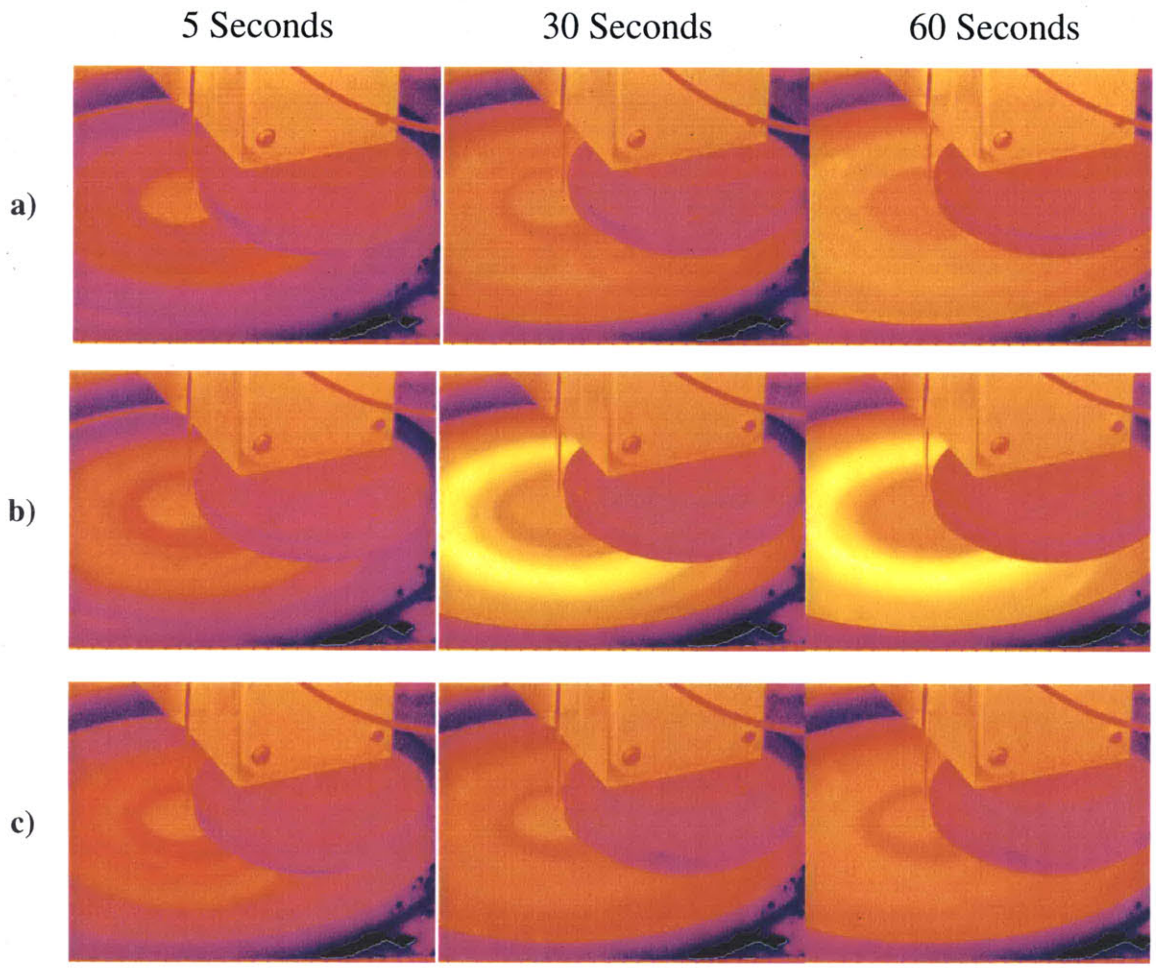

Figure 13. Snapshots of polishing: a) baseline process, b) high speed process, c) low slurry process

taken at different times during the polish of a wafer under all three recipes. In these images the lighter colors represent the higher temperature regions, and the darker colors represent the cooler regions. It is evident by looking at the images that all three processes increase in average temperature as the polishing proceeds. Another observation is that although the baseline and low slurry recipes appear to produce basically the same results, the high speed process is immediately identifiable by its higher temperature. This corresponds to the work done with the point sensor where the high speed process stood out. In the earlier work it was noted that the temperature profile for the high speed process was also more peaked than the other two. This aspect is also identifiable in the images. The high speed process clearly has a more sharply defined signature. The temperature drops off more quickly as one moves away from the center of the wafer track both towards the center of the pad and towards the edge of the pad. To get a better idea of the actual profiles for each of the processes, ten points were chosen evenly spaced across the bottom edge of the wafer carrier and the temperature data for these points was extracted. A typical signal extracted from a pixel of the camera image is shown in Figure 14. This signal was extracted from a polishing done at the baseline recipe. There is the expected exponential rise to a steady state temperature.

There is some noise in the signal, but when the high speed process is monitored in a similar way it was evident that the higher speeds were introducing more variation in the

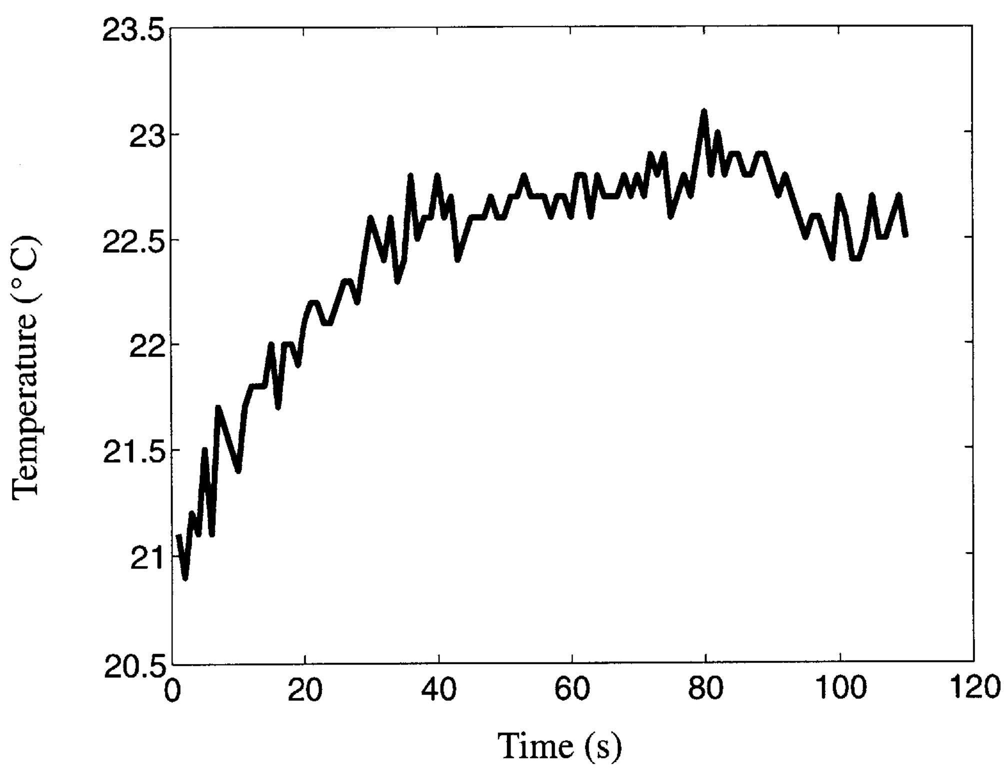

Figure 14. Temperature reading of single pixel in baseline process polish

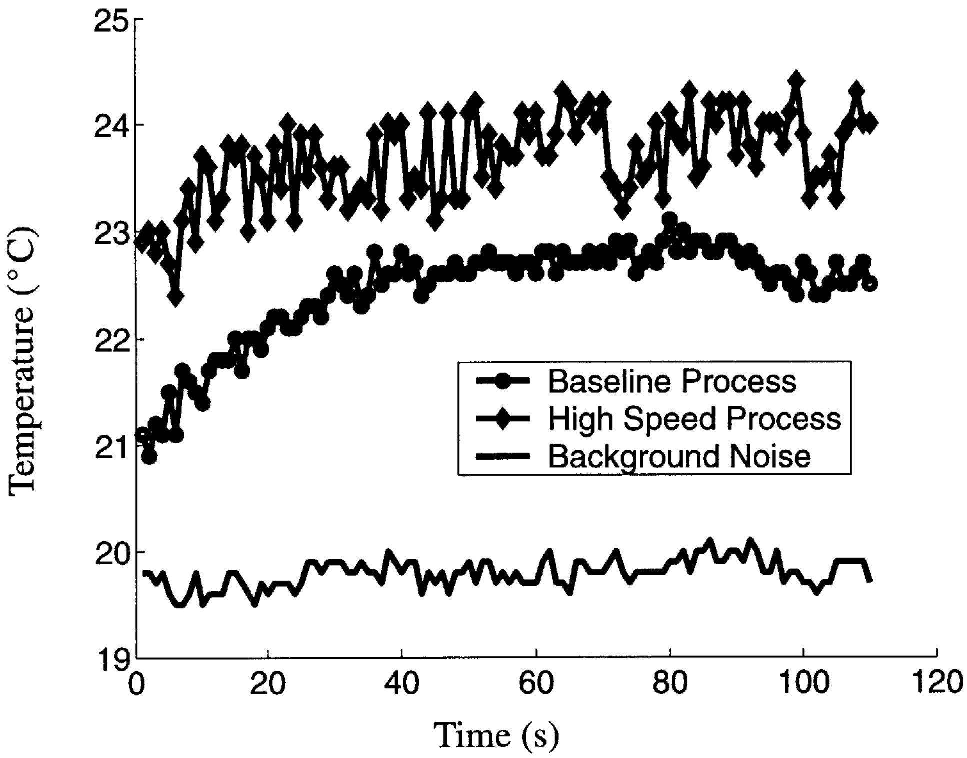

Figure 15. Temperature of single point in baseline and high speed process plotted against signal from a point of constant temperature.

temperature readings as shown in Figure 15. In this graph it is evident that the noise in the signal taken from the high speed process is much greater than the noise in the baseline process, and in the background noise which was the signal taken at a point of constant temperature during a polishing. This seems to indicate that there are other factors contributing to the noise in the high speed process. A closer look into the noise in the process found that the data taken at points close together spatially was highly correlated. This suggests that the cause is a physical one. A possible cause of the variation is the thickness of the slurry. A thicker layer of slurry would emit more infrared radiation than a thinner layer causing a higher intensity reading to be observed. In the higher speed process, a larger variation in the slurry fluid layer could cause the higher variation in the process. At this time there is no concrete explanation for this phenomenon.

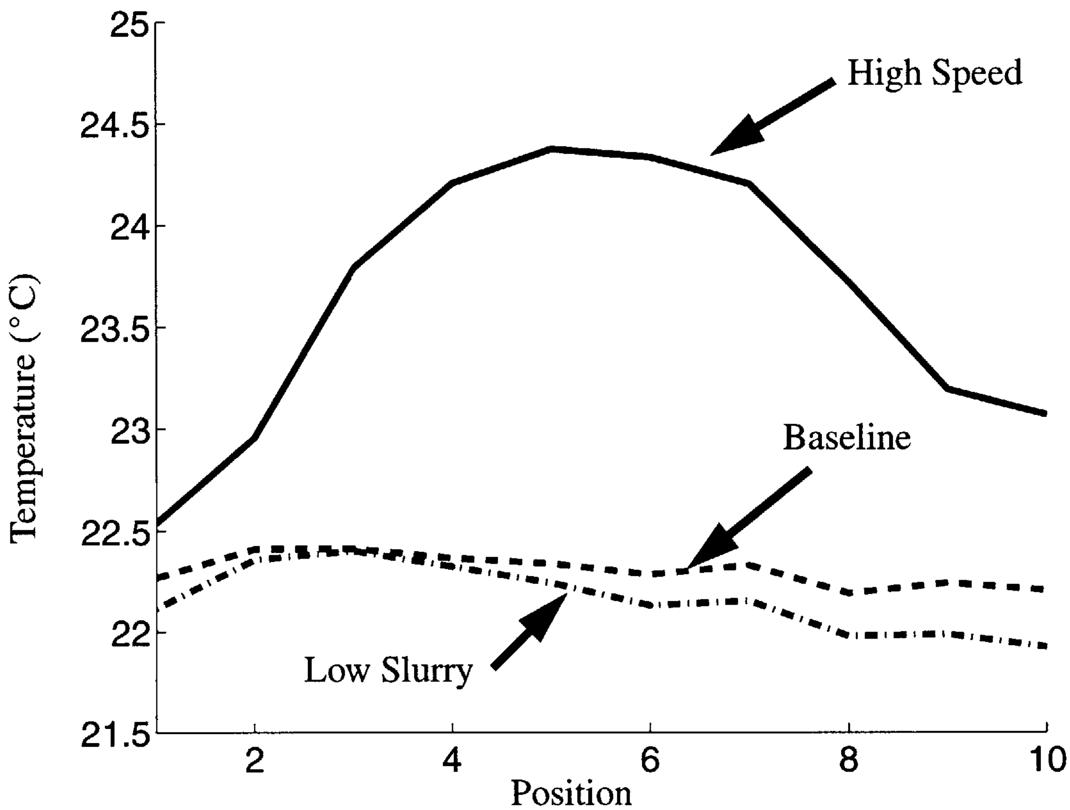

Figure 16. Temperature profiles taken from measurements at 10 points from IR camera recording

To eliminate some of the noise and transient effects, the time data for each of the ten points aligned at the edge of the wafer carrier head was averaged. Both replicates for each process were also averaged point by point. The ten points were then used to construct an average profile for each of the process settings. The average profiles are shown in Figure 16. The profiles acquired from the camera closely resemble those acquired by the point sensor. Again the baseline and low slurry processes are much closer together and have a very flat profile while the high speed process has a greater curvature and a higher average temperature. This result is encouraging, because is attests to the repeatability of the results and the significance of the relation between temperature and process settings. Such a strong relationship will lend well to modeling in the future.

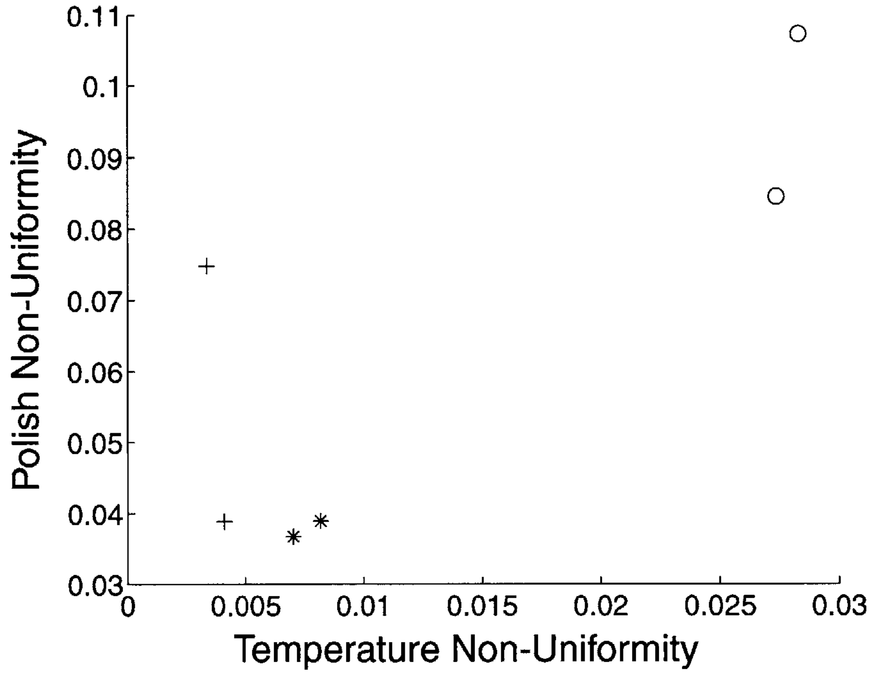

Figure 17. Comparison of temperature profile uniformity and post polish oxide thickness uniformity of wafers

The other goal of this experiment was to demonstrated the connection between the uniformity of the temperature profiles, and the uniformity of the final film thickness of the post polish wafers. This relation is described in Figure 17, where non-uniformity is defined as the standard deviation of the measurements over their mean as it was in the previous chapter. As in the experiments done with the IR point sensor, the relationship between the temperature profile non-uniformity and the final film non-uniformity appears to be linear, with the exception of one point. This point represents the measurements taken on the first wafer which was polished in the experiment. The film thickness surface map shows that this wafer, when compared to the others polished, was highly irregular: the center showed almost no signs of polishing. This pattern of variation in polish was not observed in any other wafer that was polished during the experiment. It is hypothesized that this deviation is due to the fact that no dummy wafers were polished after the setup of the tool. For this reason, the properties of the pad were different during this run than they were for the other wafers creating a deviation in readings.

In comparing Figure 17 based on simultaneous samples of 10 spatial points gathered with the IR camera during a single wafer polish, with Figure 11 where we utilized the single point sensor, we see that in both cases there exists a strong relationship between the ins-situ observations of temperature and the resulting wafer polish non-uniformity.

Chapter 4

Conclusions and Future Work

The work described in previous chapters has shown that there is great promise in the implementation of an infrared camera as a sensor for the CMP process. There are many directions that such work could take because of the thermal dependence of both chemical and mechanical reactions in the system. The obvious first step is to construct models of the dependencies observed. There is also some promise in the development of an end point detection algorithm that would enable spatial endpoint information to be extracted from the thermal signal.

4.1 Conclusions

It is clear that the temperature measurements taken by the infrared camera support the earlier findings with the infrared point sensor. In both cases, the averaged temperature profiles are similar, though the degree of detail available in the camera's case is much higher. In both experiments the high speed process generated a more pronounced temperature signature than the other two processes which corresponds with the higher removal rate observed by this process.

When the uniformity of the temperature profile was compared with that of the post polish film thickness of the wafers, again the experiments agree. Both show that there is what appears to be a linear dependence between the two variables.

4.2 Future work: Modeling

The first area of knowledge which needs to be addressed is the modeling of the infrared data with regards to process uniformity and process recipe. If an empirical or physical model can be constructed using this data it would allow in-situ measurement of wafer-level uniformity. It could also be useful to monitor the condition of the process. Despite the fact that a specific recipe is used, other variables can drift in the process such as the pad condition. The infrared camera would be able to provide information on this drift while sacrificing the least number of wafers. A proper design of experiments varying major process settings would provide a great deal of information on the dependencies between the process variables and the infrared signature observed by the camera during polishing.

Another effect which would be beneficial to characterize is the pad break-in period. It was evident that in the first experiment, this phenomenon was captured in the temperature of the polishing wafers, probably due to the relation between the removal rate and the observed temperature. If the infrared sensor could be used to monitor the pad break in, it would minimize the number of dummy wafers necessary to bring the pad to a reliable and usable condition.

4.3 Future work: Metal Endpoint Detection

Currently infrared sensors are utilized to detect endpoint in metal polishing processes. When one point is observed for the duration of a polish the endpoint can be detected either by the sudden rise or fall of the temperature signal. To look into similar applications for the infrared camera, a copper wafer was polished. During the polishing the temperature of

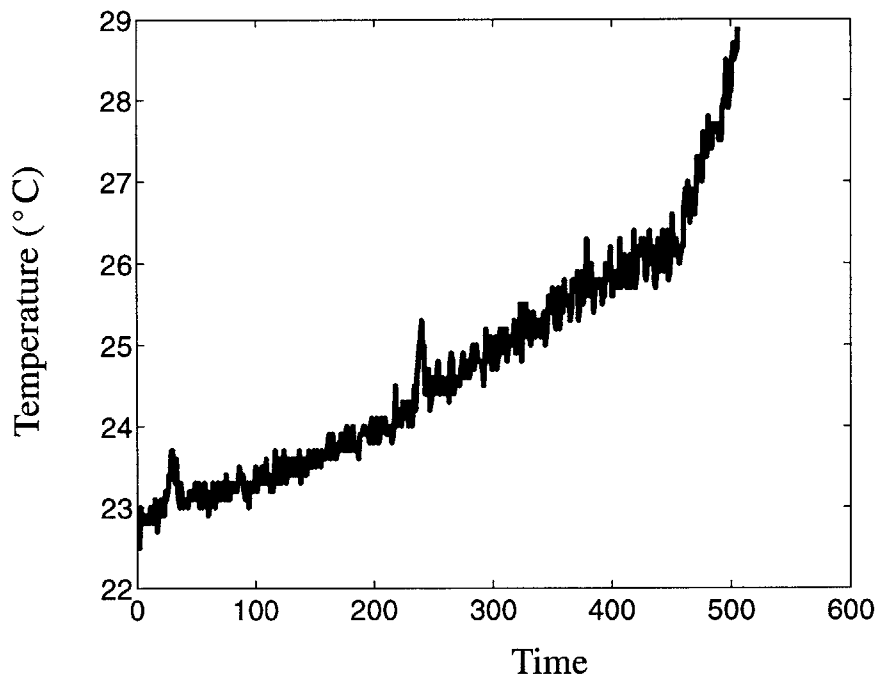

Figure 18. One point time trace of temperature during copper polishing

the wafer was recorded with the infrared camera. The temperature during the polish for one point is shown in Figure 18. This provides a general idea of when the endpoint occurs, by identifying the slope change in the temperature trace. The benefit of having multiple sensors, as in the case of an infrared camera, is the possibility of spatial EPD. The temperature can be monitored for many radii of the wafer, so rather than knowing only when one radii or some average area has endpointed, it can be determined how uniformly the wafer is reaching endpoint by monitoring the time of endpoint for many pixels in the camera picture. To look into this, we observed many points along the edge of the carrier head. The temperature profiles in Figure 19 are taken from the same wafer at different times during the polishing. The starting wafer was a blanket copper wafer with one micron of copper deposited on the surface. This was polished off and the temperature was recorded by the infrared camera. The vertical axis in the figure is the measured temperature, and the horizontal axis of each profile represents the spatial position of the pixel monitored. Five pixels wee used for this analysis. They were evenly spaced across the polishing track of the wafer, with point 1 closest to the center of the polishing pad. Each profile represented in the graph is an average over ten seconds at the given interval. This averaging was done to reduce the influence of random noise in the shape of profile. As you can see in the figure, the average temperature of the pad increased approximately linearly from the beginning of polish to 500 seconds into the polish. Between 500 seconds, and 600 seconds however, there is a drastic change in the profile observed. The temperature shoots up at a rate clearly greater than that previously observed, and the profile

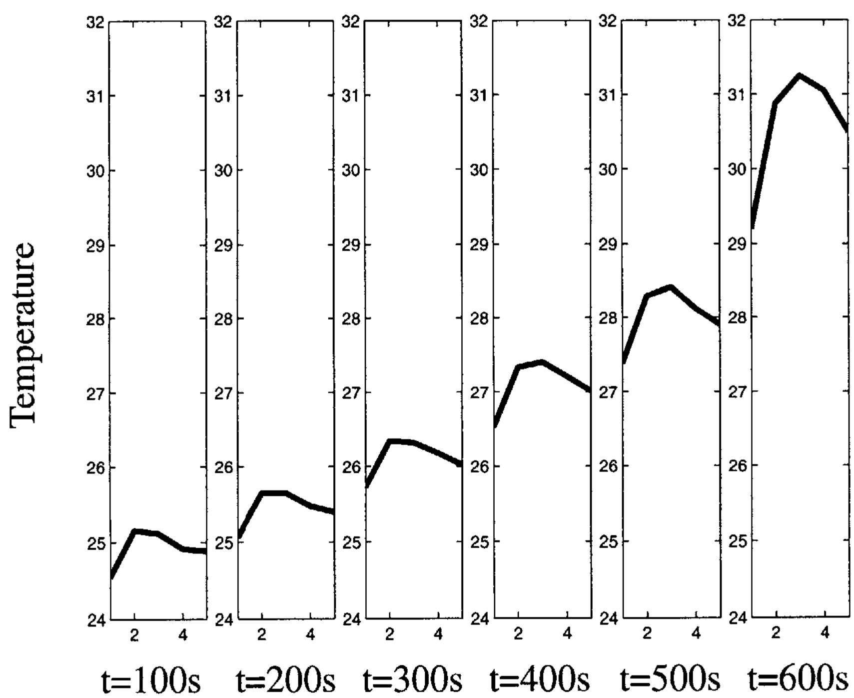

Figure 19. Averaged profiles of copper CMP process at 100s spaced intervals

changes dramatically. This is where we believe endpoint occurred for a portion of the wafer. The profile itself is also changing in character. It slowly becomes more pointed in the center, and between 500 and 600 seconds, there is a clear peaking in the shape of the profile. These results seem promising, however, because of the non-uniformity of the finished wafer, we were unable to make quantitative conclusions about the endpoint detection. The dramatic change in temperature is encouraging for future studies in spatial EPD using the infrared camera.

References

[1] J. Steigerwald, S. Murarka, R. Gutman, *Chemical Mechanical Planarization of Micro-electronic Materials*, 1997.

[2] J. Steigerwald, S. Murarka, R. Gutman, *Chemical Mechanical Planarization of Micro-electronic Materials*, 1997.

[3] B. Stine, D. Ouma, R. Divecha, D. Boning, J. Chung, D. Hetherington, I. Ali, G. Shinn, J. Clark, O.S. Nakagawa, S.-Y. Oh, *A closed-form analytic model for ILD thickness variation in CMP processes,* Proceedings of 2nd Intn'l CMP-MIC, pp 266-273, 1997.

[4] T. Bibby, K. Holland, *In Situ Endpoint Detection for CMP,* Talk given at meeting of CMP Users Group of N.California chapter of the AVS, 1998.

[5] H.-W. Chiou, L.-J. Chen, H.-C. Chen, *On monitoring CMP removal rate by in-situ temperature measurements*, Proceedings of 2nd Intn'l CMP-MIC, pp 131-138, 1997.

[6] R.L. Lane and G. Mylnar, *Real Time Infrared Imaging of Pads During the CMP Process*, Proceedings of 2nd Intn'l CMP-MIC, pp 139-142, 1997.