



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

ارائه روشی نوین برای تطبیق استریو به کمک مجموع فاصله همینگ

سکینه شیرازی تهرانی و پیمان معلم و محسن عشوریان

دانشگاه آزاد اسلامی واحد نجف آباد, دانشکده برق Fars423@yahoo.com

دانشگاه اصفهان, دانشکده فنی مهندسی rp_moallem@eng.ui.ac.ir

دانشگاه آزاد اسلامی واحد مجلسی, دانشکده برق mohsena@yahoo.com

چکیده

بینایی استریویا داشتن حداقل دو تصویر از یک صحنه، می توان به استخراج عمق پرداخت. در این راستا مهمترین کار افزایش سرعت اجرا و دقت و در پی آن کاهش پیچیدگی محاسباتی است. هر چقدر بتوان فضای جستجو را کاهش داد، نه تنها مرحله تطابق سریعتر اجرا می شود. بطور کلی الگوریتم های مختلف بینایی استریو اعم از الگوریتم های بر مبنای پیکسل، بر مبنای سطح، و بر مبنای ویژگی می باشند که با توجه به سرعت اجرا، ترکیب روش مبتنی بر ویژگی و سطح در این تحقیق انتخاب شده است. برای کاهش فضای جستجو از قیود محدودیت خط مبنا استفاده شده است. جهت افزایش دقت از معیار شباهت مجموع فاصله همینگ و برای افزایش سرعت از دو مرحله برنامه نویسی پویا استفاده شده است.

واژه های کلیدی: تطابق استریو. بینایی استریو. خط $epipolar$. برنامه نویسی پویا. مجموع فاصله همینگ

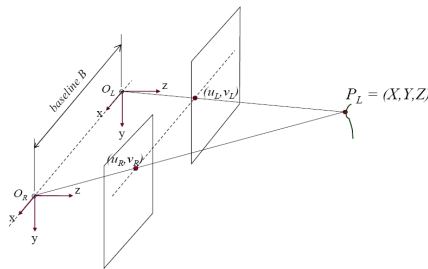


بینایی استریو

در این قسمت به هندسه سیستم استریو و الگوریتم های تطابق استریو و بیان معیارهای شباهت پرداخته می شود [3,2,7].

هندسه سیستم استریو

همان طور که پیشتر نیز ذکر شد، برای داشتن تصاویر سه بعدی صحنه باید دو تصویر مجزا از دو نقطه که فاصله اندکی دارند تهیه کرد بنابراین از دو دوربین مشابه که با فاصله کمی از یکدیگر قرار دارند استفاده می کنیم. در شکل (1-2) نمونه ای از یک سیستم استریو و نحوه ی قرار گیری دوربین ها نشان داده شده است. O_R و O_L جایگاه دو دوربین استریو با فواصل کانونی یکسان f ، در صحنه را نشان می دهد. فاصله ی بین این دو نقطه، در سیستم های استریوی گوناگون، متفاوت است و تحت عنوان خط اپی پولار شناخته می شود.



شکل (1-2). هندسه ی سیستم استریوی مبنا

در شکل (1-2)، نقطه ی P_L نقطه ای در فضای سه بعدی است و (u_L, v_L) و (u_R, v_R) تصاویر تشکیل شده نقطه P_L در دوربین های چپ و راست است. در این شکل صفحات تصویر دو دوربین در راستای یکدیگرند و O_L و O_R مرکزهای مختصات دو دوربین چپ و راست هستند.

در یک سیستم استریو، دو تصویر از دو زاویه ی مختلف یک صحنه گرفته می شود و سپس با ترکیب این دو، تصاویر سه بعدی، صحنه بدست می آید. پیاده سازی استریو در طی سه مرحله انجام می شود: کالیبراسیون دوربین، تعیین نقاط ویژه دو تصویر و تطابق استریو برای یافتن مختصات سه بعدی.، مهمترین مسئله برای تولید تصاویر سه بعدی، مسئله ی تطابق استریو است. در تطابق استریو، دو نقطه ی متناظر در دو تصویر، باید نمایانگر نقطه ی فیزیکی یکسانی در صحنه ی واقعی باشند [4].

در این راستا مهمترین کار، افزایش سرعت اجرا و دقت و همچنین کاهش پیچیدگی محاسباتی است. کاهش فضای جستجو، به طور خطی در افزایش سرعت اجرای مرحله ی جستجو موثر می باشد. بنابراین استفاده از محدودیت خط اپی پولار باعث کاهش زمان اجرا و خطای تطابق می گردد [5].

در حالت کلی الگوریتم های تطابق استریو به سه دسته تقسیم می شوند: الگوریتم مبتنی بر ناحیه، الگوریتم مبتنی بر مشخصه، الگوریتم مبتنی بر پیکسل و روش های ترکیبی [1].

در این مقاله با توجه به اهمیت سرعت اجرا، از میان الگوریتم های مختلف بینایی استریو، الگوریتم مبتنی بر ناحیه و مبتنی بر مشخصه استفاده کرده ایم. معیارهای مختلف شباهت مورد استفاده و با هم مقایسه کرده ایم. برای کاهش فضای جستجو از قیود مختلفی استفاده کرده ایم و نیز برای افزایش سرعت اجرا از دو مرحله برنامه نویسی پویا بهره برده ایم.



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

زده می شود. این روش برای تصاویر شلوغ بسیار مفید است. از جمله معروفترین این الگوریتمها، الگوریتم برنامه نویسی پویا را می توان نام برد.

در الگوریتمهای بر مبنای ویژگی، در ابتدا از هر دو تصویر چپ و راست عناصر مورد تطابق استخراج می شوند. سپس در مرحله تطابق، شباهت آنها سنجیده شده و عناصر متناظر مشخص می گردند. ایجاد تطابق در روشهای بر مبنای ویژگی سریعتر از تطابق بر مبنای سطح است. در این روش اختلاف مکانی مربوط به نقاط ویژه را به می دهد. بنابراین برای تمام نقاط باید از ترکیب با روش مبتنی بر سطح استفاده کرد.

فیود سیستم استریو

محدودیت خط ایپی پولار (epipolar): همان طور که گفتیم، در هندسه ی تصویر برداری استریو، مقدار Y در صفحه ی دو دوربین چپ و راست یکسان است. بنابراین اگر یک نقطه در صفحه ی دوربین چپ با مختصات (x_l, y_l) مشخص باشد، برای یافتن متناظر آن در صفحه ی دوربین راست، باید روی خط $y = y_l$ جستجو انجام گیرد. به عبارت دیگر به جای همه ی صفحه ی دو بُعدی، تنها در یک فضای یک بُعدی جستجو انجام می شود. کاهش فضای جستجو از کل تصویر دو بُعدی به یک خط را قید یا محدودیت خط ایپی پولار می گویند.

محدودیت عمق: در دنیای واقعی همواره عمق دارای محدوده ی تعیین شده ای است و از حدود خاصی تجاوز نمی کند:

$$Z_{\min} < Z < Z_{\max} \quad (3-1)$$

با توجه به رابطه ی (3-1) می توان برای اختلاف مکانی نیز به صورت زیر محدوده ای تعریف کرد.

$$(u_L, v_L) = \left(f \frac{X}{Z}, f \frac{Y}{Z} \right)$$

$$(u_R, v_R) = \left(f \frac{X - B}{Z}, f \frac{Y}{Z} \right) \quad (1-1)$$

طبق معادله ی (1-1)، مقدار v برای دو تصویر چپ و راست یکسان می باشد. برای بدست آوردن مقدار اختلاف مکانی می توان از معادله ی (2-1) استفاده کرد.

$$u_L - u_R = f \frac{B}{Z} \quad (2-1)$$

مقدار $u_L - u_R$ اختلاف مکانی نامیده شده و با d نمایش داده می شود. می توان عمق هر نقطه را با توجه به مقدار اختلاف مکانی آن با توجه به معادله ی (3-1) بدست آورد:

$$d = u_L - u_R = f \frac{B}{Z} \Rightarrow Z = f \frac{B}{d} \quad (3-1)$$

بنابراین عمق نقطه متناسب با عکس اختلاف مکانی می باشد. در این روابط f فاصله کانونی لنز دوربینها و B خط ایپی پولار است. [3]

الگوریتم های بینایی استریو

در حالت کلی، الگوریتمهای تطبیق استریو به سه گونه ی بر مبنای پیکسل، بر مبنای سطح و بر مبنای ویژگی تقسیم می شوند. در روشهای بر مبنای پیکسل، سعی می شود از ابتدا مقدار اختلاف مکانی برای تمامی پیکسلها محاسبه شود. این روش نسبت به دو روش بعدی، زمان اجرای طولانی تری دارد و لذا برای کاربردهایی که سرعت اجرا در آنها مهم است، مناسب نمی باشد.

در روش الگوریتم بر مبنای سطح، بلاکی در نظر می گرفته و میزان شباهت بین دو بلاک در دو تصویر استریو تخمین



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

جمع قدر مطلق تفاضل‌ها (SAD): این معیار همان SSD است که به جای توان دو از قدر مطلق استفاده کرده است.

$$\begin{aligned} d_{\min} &< d < d_{\max} \\ d_{\min} &= b \cdot f / Z_{\max} \\ d_{\max} &= b \cdot f / Z_{\min} \end{aligned} \quad (3-1)$$

$$SAD = \sum_w |w_{i,j} - w'_{i,j}|$$

ضریب همبستگی متقابل نرمال (NCC): در این رابطه، \bar{w} و \bar{w}' مقادیر متوسط در پنجره‌های w و w' می‌باشند. استفاده از NCC محاسبات نسبتاً زیادی دارد.

$$NCC = \frac{(w - \bar{w})(w' - \bar{w}')}{|w - \bar{w}| |w' - \bar{w}'|}$$

مجموع فاصله همینگ (SHD):

$$SHD = \sum_{i,j} \text{bitwiseXOR } I_2(x+i, y+j)$$

استفاده شده است. خطای SHD در این مقاله از روش دقت بیشتر است. این روش کمتر و

برنامه نویسی پویا:

کل (3-1) قسمت (الف)، شبکه‌ای را نشان می‌دهد که برای یافتن تطابق نقاط روی دو خط اسکن از دو تصویر چپ و راست، باید جستجو شود. در این شکل هر یک از دو خط اسکن دارای 10 پیکسل هستند و ماکزیموم اختلاف مکانی که با d_{\max} نشان می‌دهیم، نیز در این مثال 3 پیکسل در نظر گرفته شده است. هر سلول (x, y) در این شبکه به معنای یک تطابق ممکن بین پیکسل x در تصویر سمت چپ و پیکسل y در تصویر سمت راست است. الگوریتم مطرح شده بهترین مسیر ممکن که از ستون اول تا آخرین سطر گسترده شده را جستجو می‌کند. در این شکل، پیکسل‌های تطابق یافته با M نشان داده شده‌اند. همان طور که شکل (3-1) قسمت (ب) نشان می‌دهد، بین هر جفت پیکسل سه حرکت مجاز

محدودیت ترتیب: برای دو جفت نقطه‌ی متناظر، ترتیب چپ و راست بودن آن دو نقطه در یک تصویر به همان صورت در تصویر دیگر رعایت شود.

معیارهای تطابق استریو:

معیارهای گوناگونی برای بررسی شباهت میان نواحی،

وجود دارند که عمومی‌ترین آنها روابط بین سطوح

روشنایی پنجره‌ی چپ و راست می‌باشد. در اینجا به سه نمونه اشاره می‌کنیم.

جمع مربع تفاضل‌ها (SSD): مقدار این معیار با شباهت

دو پنجره از دو تصویر چپ و راست، رابطه‌ی عکس دارد:

یعنی هر چه شباهت بیشتر باشد، مقدار SSD کمتر

خواهد بود. برای استفاده از این معیار باید یک مقدار

آستانه تعیین شود. اگر مقدار SSD برای دو نقطه‌ی خاص

از حد آستانه کمتر باشد، آن دو نقطه مطابق یکدیگر

فرض می‌شوند. این معیار بار محاسباتی کمی دارد. در

معادله‌ی SSD که در ادامه آمده است، w و w' دو پنجره

مستطیلی از تصاویر چپ و راست هستند. در این معادله

توان دو وجود دارد که متناظر با ضرب است. برای حذف

این ضرب و راحتی پیاده سازی آن، می‌توان از معیار

ASD استفاده کرد.

$$SSD = \sum_w (w_{i,j} - w'_{i,j})^2$$



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

(1-3)

وجود دارد. دایره‌ها در این تصویر، خانه‌های شبکه را نشان می‌دهند.

$$C(x, y) = D(x, y) + m$$

$$m = \min \left\{ \begin{array}{l} C(x-1, y-1), \\ C(x-2, y-1) + k_{occ}, C(x-3, y-1) + k_{occ}, \dots, \text{until } x = y \\ C(x-1, y-2) + k_{occ}, \dots, C(d_{\text{Max}}, x - d_{\text{Max}} - 1) + k_{occ} \end{array} \right.$$

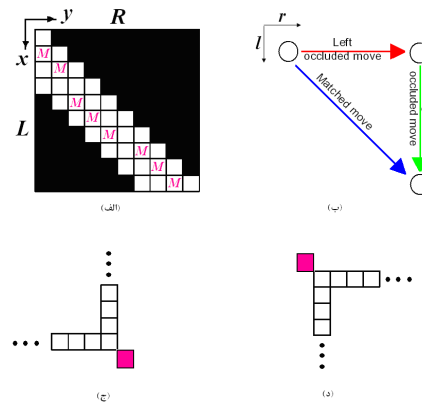
تطابق‌های بلافاصله قبلی و بلافاصله بعدی هر تطابق (x_i, y_i) در قسمت‌های (ج) و (د) به ترتیب نشان داده شده‌اند. با توجه به قسمت (الف) این شکل، هر تطابق $d_{\text{max}} + 1$ کاندید ممکن برای تطابق‌های بلافاصله قبلی و $d_{\text{max}} + 1$ کاندید ممکن برای تطابق‌های بلافاصله بعدی دارد.

در معادله (1-3)، K_{occ} ثابت جریمه‌ی گم شدگی و $D(x, y)$ مقدار شباهت بین پیکسل‌های x و y است. وقتی مقادیر ماتریس C همگی محاسبه شدند، سلولی با کمترین هزینه از سطر آخر این ماتریس به عنوان آخرین تطابق انتخاب می‌شود. سپس با شروع از این سلول، ماتریس P برای یافتن دنباله‌ی بهینه جستجو می‌شود.

برای هر سلول سفید در شکل (1-3) قسمت (الف)، مقدار $C(x, y)$ را به عنوان هزینه‌ی بهترین دنباله تطابق تا نقطه‌ی (x, y) و $P(x, y)$ را به عنوان نشانه‌ی گری به تطابق بلافاصله قبلی، ثبت می‌کنیم. در هر سلول سفید در شکل (الف)، مقدار شباهت بین آن دو پیکسل را قرار می‌دهیم. ما این ماتریس را M می‌نامیم و سپس این ماتریس را نرمال می‌کنیم. مقادیر نرمال شده‌ی ماتریس M را در ماتریسی به نام N قرار می‌دهیم.

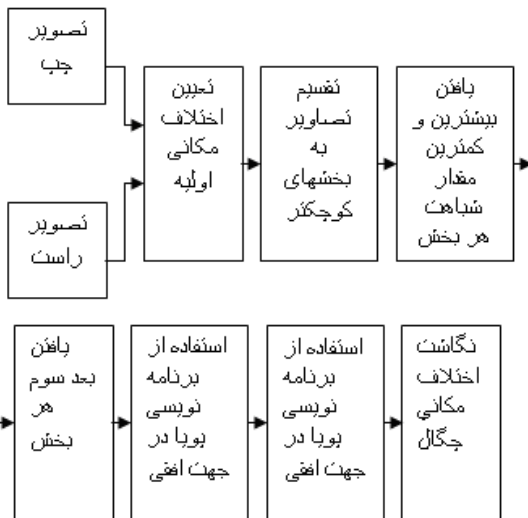
4- روش پیشنهادی

در این بخش الگوریتم مورد نظر مطرح می‌شود. همان طور که قبلاً اشاره شد، الگوریتم مبتنی بر ویژگی لبه و سطح انتخاب شده است.



شکل (1-3). (الف). شبکه‌ی جستجو و یک دنباله‌ی تطابق (سلول‌های M), (ب). سه حرکت مجاز بین هر دو پیکسل در شبکه. (ج). تطابق‌های بلافاصله قبلی (د). تطابق‌های بلافاصله بعدی

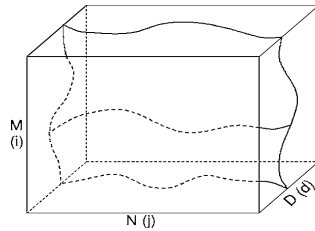
برای محاسبه‌ی هزینه‌ی بهترین مسیر به هر سلول در این شبکه، ما از فرمول (1-3) استفاده می‌کنیم:



شکل (1-4) - بلوک دیاگرام الگوریتم پیشنهادی



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388
جهت عمودی انجام داده بهترین مقدار چگالی را در
ماتریس سه بعدی برای هر پیکسل انتخاب می کنیم.



شکل (4-4) - یافتن بیشترین مقدار اختلاف مکانی بروش
برنامه نویسی پویا

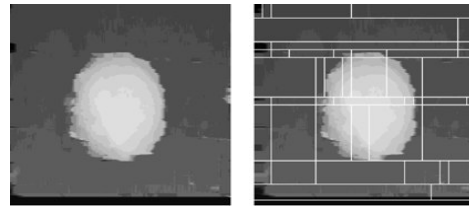
خروجی الگوریتم، ماتریس اختلاف مکانی چگال (Disparity Map) است که در هر خانه ی این ماتریس،
عمق متناظر با آن مختصات قرار دارد.

5- نتیجه گیری:

بدلیل اطلاعات زیادی که در عمق تصاویر ذخیره شده
است بینایی سه بعدی اهمیت خیلی زیادی پیدا می کند.
لذا الگوریتم های زیادی در زمینه بینایی استریو مطرح
شده اند که هر یک محدودیت ها و شرایط خاص خود را
دارند و هنوز هیچ یک به سرعت و دقت مورد نیاز برای
استفاده در جهان واقع، دست نیافته اند. بنابراین ارائه
روشهایی برای بهبود این الگوریتم ها گام ارزشمندی در
جهت حل مسائل بینایی سه بعدی خواهد بود. در اینجا ما
از معیار شباهت SHD استفاده کرده ایم که باعث
افزایش دقت شده است.

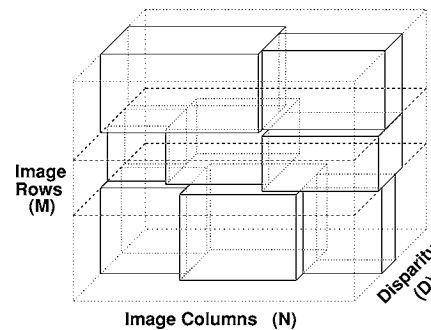
نتایج الگوریتم بر روی یک تصویر نمونه در شکل 5-1
آمده است. مشاهده می کنیم که دقت در تصویری که از
معیار SHD استفاده شده بیشتر از تصاویر دیگر است.

در این مقاله در ابتدا بر اساس روش الگوریتم ویژگی لبه
ها اختلاف مکانی اولیه را بدست آورده سپس بر اساس
آن با توجه به مقدار هر پیکسل با مقادیر پیکسلهای
همسایگی شکل به بخشهای کوچکتر جهت افزایش
سرعت تقسیم می گردد. شکل زیر نمونه ای از تقسیم
تصویر را نشان می دهد.



شکل (2-4) - تقسیم تصویر به بخشهای کوچکتر

پس بیشترین و کمترین مقدار ها اختلاف مکانی هر
قسمت را یافتیم. پنجره ای $7*7$ در نظر گرفته روی
تصویر راست شیفتمی دهیم مقدار d را از کمترین تا
بیشترین مقدار تغییر داده برای هر بخش یک ماتریس سه
بعدی یافت می گردد. شکل زیر نمای ماتریس را نشان
می دهد.

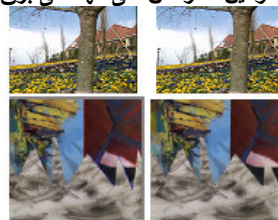


شکل (3-4) - نمای سه بعدی مقادیر بدست آمده حاصل از
شیفتم پنجره شباهت

بعد از این مرحله برای یافتن تطابق بین تصاویر یک
مرحله برنامه نویسی پویا در جهت افقی و یک مرحله در

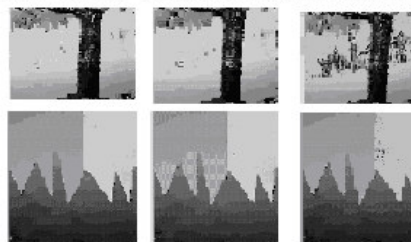


دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388



تصویر چپ تصویر راست

SHD Disparity Map NCC Disparity Map SAD Disparity Map



شکل (5-1) - نتایج پیاده سازی با معیار شباهت های

مختلف

مراجع :

- [1] Camillo J. Taylor, "Surface Reconstruction from Feature Based Stereo", Proceedings of the Ninth IEEE International Conference on Computer Vision (ICCV'03) 0-7695-1950-4/03 \$ 17.00 © 2003 IEEE.
- [2] Ken-Ichiro Tsutsui, Masato Taira, Hideo Sakata, "Neural mechanisms of three-dimensional vision", Neuroscience Research 51 (2005) 221-229.
- [3] Andreas Koschan, "What is New in Computational Stereo Since 1989: A Survey on Current Stereo Papers", Technische Universität Berlin, Technischer Bericht 93-22, August 1993.
- [4] P. Moallem and K. Faez, "Effective Parameters in Search Space Reduction Used in a Fast Edge-Based Stereo Matching", Journal of Circuits, Systems, and Computers, Vol. 14, No. 2(2005) 249-266 © World Scientific Publishing Company, Received 23 November 2002, Revised 11 June 2004.
- [5] P. Moallem, M. Ashorian, B. Mirzaeian, M. Ataei, "A Novel Fast Feature Based Stereo Matching Algorithm with Low Invalid Matching", WSEAS Transaction on Computers, Issue 3, Vol. 5, March 2006.
- [6] هاجر صادقی، سید امیر حسن منجمی و پیمان معلم، "ارائه روشی نوین برای تطبیق استریو در تصاویر رنگی به کمک زنجیره های پیوسته"، سومین کنفرانس دانشگاه فردوسی مشهد، آذر ماه سال 1386
- [7] Changming Sun, "Fast Stereo Matching Using Rectangular Subregioning and 3D Maximum-Surface Techniques", International Journal of Computer Vision 47(1/2/3), 99-117, 2002



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388

Fast Method in Stereo Matching using Som of Hamming Distances

Sakineh shirazi tehrani-peyman moallem-mohsen ashorian
Azad University of Najaf Abad
Fars423@yahoo.com

University of Isfahan
p_moallem@eng.ui.ac.ir

Azad University of Majlesi
mohsena@yahoo.com

Abstract

Depth extraction is one the most popular application of stereo vision. Reducing the computational complexity of the correspondence stage which is the most complex stage in the stereo algorithms, can effectively decrease the execution time. The correspondence is a search problem. Therefore, search space reduction which is the main discussion of this thesis, can decrease the execution time as well. Matching techniques can be divided broadly into area-based and feature-based or a combination of them and pixel-base image matching,.

In this paper we employ combination of them stereo matching as fast and reliable methods, where some matching constraints such as epipolar line. we use correlation Som of Hamming Distances , two-stage dynamic programming (TSDP) as fast and reliable methods.



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388



دومین کنفرانس ملی مهندسی برق ایران - دانشگاه آزاد اسلامی واحد نجف آباد - 5 و 6 اسفند ماه 1388