Medical Image Processing for Diagnostic Applications

MRI – Acquisition Devices

Online Course – Unit 19
Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

MR Acquisition Devices

Summary

- Take Home Messages
- Further Readings

MR Acquisition Devices

MR scanners are huge and heavy systems with strong superconducting magnets:

Figure 1: Examples of 3 Tesla systems: MAGNETOM Verio (left) and Trio Scanner (right) (images courtesy of Siemens Medical Solutions)

Physical and Mathematical Ingredients

For MRI imaging, the following physical and mathematical concepts are required to be understood:

- **nuclei** serve as objects to be imaged,
- **homogeneous magnetic fields** are generated by the scanner to align the nuclear moment vectors,
- the **resonance phenomenon** that results from the interaction of nuclei with the magnetic field enables measurements,
- **Fourier methods** are used for image reconstruction,
- **image enhancement algorithms** are applied to compensate for violations of the required homogeneity of the magnetic field (**bias field correction**).

Note: Details in physics are not in the focus of this course, but the algorithmic aspects.

Components of an MR Scanner

The major four components of an MR Scanner are:

- the main magnet,
- a magnetic field gradient system,
- a radio frequency system (RF system),
- and the imaging system.

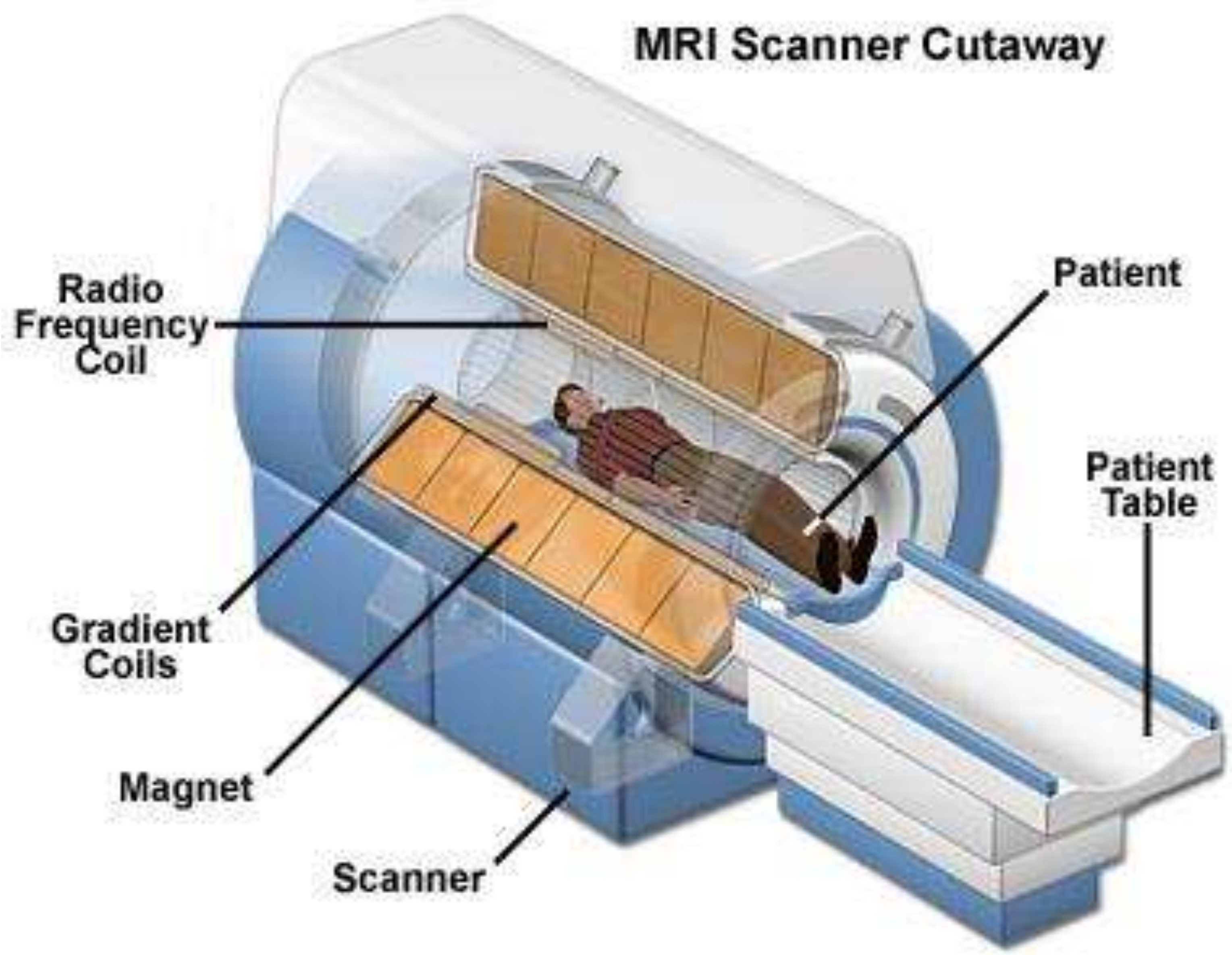

Figure 2: Main components of an MR scanner (image courtesy of the National High Magnetic Field Laboratory, Florida)

Components of an MR Scanner: Main Magnet

The **main magnet** is required to generate a strong uniform static magnetic field for the polarization of nuclear spins.

In practice, there are several options:

- **permanent magnet**: low field applications (< 0.3 T),
- **resistive magnet**: low field applications (< 1.5 T),
- **superconducting magnet**: used for higher magnetic field strengths (high end research scanners, whole-body up to 11.75 T, general research > 20 T).

The main magnet of the MAGNETOM Verio scanner is only 6.5 tons!

Necessity of Gradients

To distinguish between the nuclei, the idea is to associate a unique magnetic field with each type of nucleus. This can be achieved by continuous variation of the magnetic field dependent on the 3-D position of the nuclei.

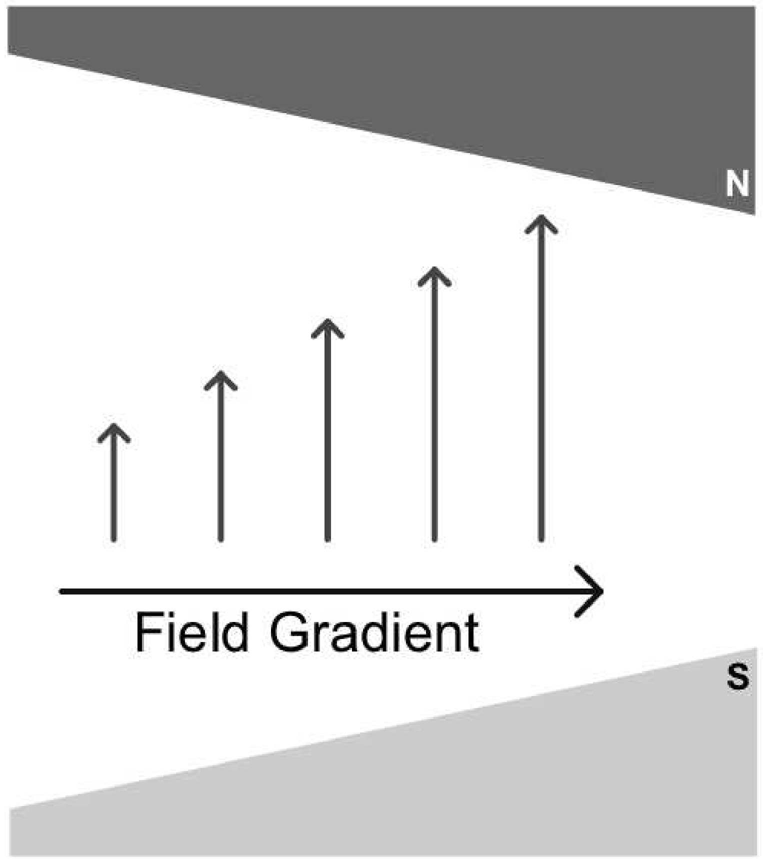

Figure 3: Schematic of continuously changing the magnetic field in 3-D

Note: The gradient strength in 3-D can be used for motion compensation. If the patient motion is known, the gradient field can be adjusted properly!

Components of an MR Scanner: Gradient System

The **magnetic field gradient system** is required to generate magnetic fields of well-defined and controlled spatial inhomogeneity as a function of the particular x-, y- and z-coordinates in space.
The gradient field is needed for signal localization in space.

- Gradient strength: e. g., 45 mT/m (millitesla per meter) in the Tim Trio 3T
- Rise time to ramp up gradient decides on quality of gradient: the smaller the rise time, the better (rise time today < 1.0 ms).

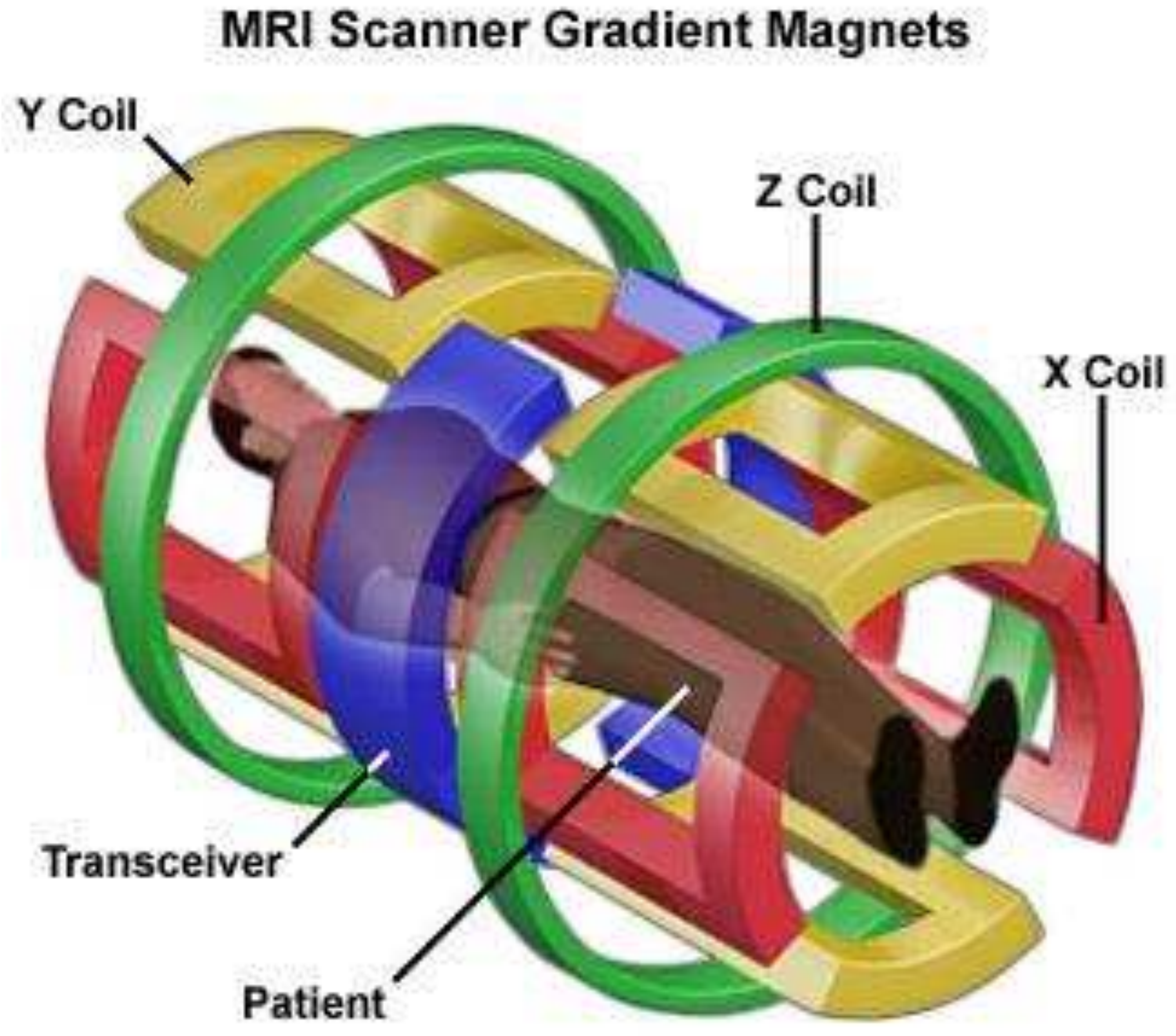

Figure 4: Principle Structure of the gradient system (image courtesy of the National High Magnetic Field Laboratory, Florida)

Components of an MR Scanner: RF System

The **radio frequency system** has two components:

- the **transmitter coil** generates a rotating magnetic field for the excitation of a spin system,
- the **receiver coil** converts magnetic changes into electrical signals.

In some systems transmission and receiver coils are identical which is then called a **transceiver coil**.

Figure 5: Examples of RF head and body matrix coils (image Siemens Medical Solutions)

Topics

MR Acquisition Devices

Summary

- Take Home Messages
- Further Readings

Take Home Messages

- Magnetic resonance imaging (MRI) requires large acquisition devices which consist of a magnet, a gradient system, an RF system, and the imaging system.
- Driven by concepts from quantum physics, a real magnetic field is not perfectly homogeneous which has to be compensated for by image processing.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

MRI – Physical Background

Online Course – Unit 20

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

On the Physics of MRI

Pros and Cons of MRI

Summary

Take Home Messages

Further Readings

Nuclear Magnetic Moments

Wolfgang Pauli (1900–1958) found out that, besides mass and charge, particles have another fundamental property, the so-called **spin**:

- The spin angular momentum is denoted by $\boldsymbol{J} \in \mathbb{R}^3$.
- A nuclear spin is associated with a microscopic magnetic field, because nuclei, like protons, carry electrical charges and rotate around their own axes if the spin is nonzero.
- The nuclear magnetic moment (or dipole moment) is denoted by $\boldsymbol{\mu} \in \mathbb{R}^3$.
- The spin angular momentum and the magnetic moment are collinear, i. e.,

$$\boldsymbol{\mu} = \gamma \cdot \boldsymbol{J}$$

  where γ is the nucleus dependent gyromagnetic ratio.

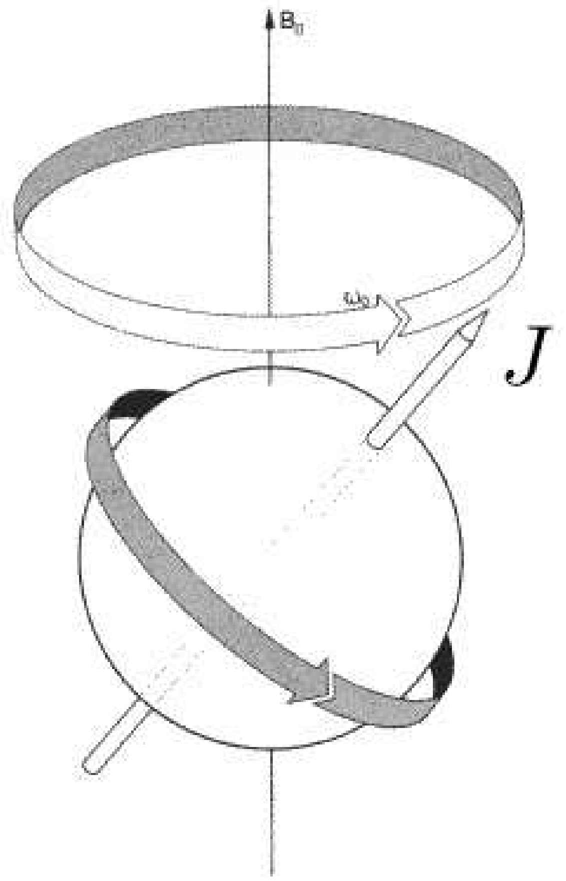

Figure 1: Scheme of the nuclear spin

Nuclear Magnetic Moments

- From the theory of quantum mechanics we know that the magnitude of the magnetic moment is given by

$$||\boldsymbol{\mu}|| = \gamma \frac{\hbar}{2\pi} \sqrt{I(I+1)},$$

  where
  - $\hbar = 6.626\,068 \times 10^{-34}\,\mathrm{m^2 kg/s}$ is Planck's constant, and
  - $I = 0, \pm\frac{1}{2}, \pm 1, \pm\frac{3}{2}, \dots$ is the spin quantum number.
- To achieve macroscopic magnetism it is required to line up spin vectors, and this can be enforced by a strong external magnetic field $\boldsymbol{B}_0$.
- The precession frequency ω_0 of $\boldsymbol{\mu}$ given by the Larmor equation $B_0 = ||\boldsymbol{B}_0||$ is

$$\omega_0 = \gamma B_0.$$

  The frequency ω_0 is called Larmor frequency.

Intensities in MRI

Intensities in MR images depend on:

- spin density,
- spin-lattice relaxation time (T1),
- spin-spin relaxation time (T2),
- molecular motion (diffusion, perfusion, flow).

MR Image Acquisition

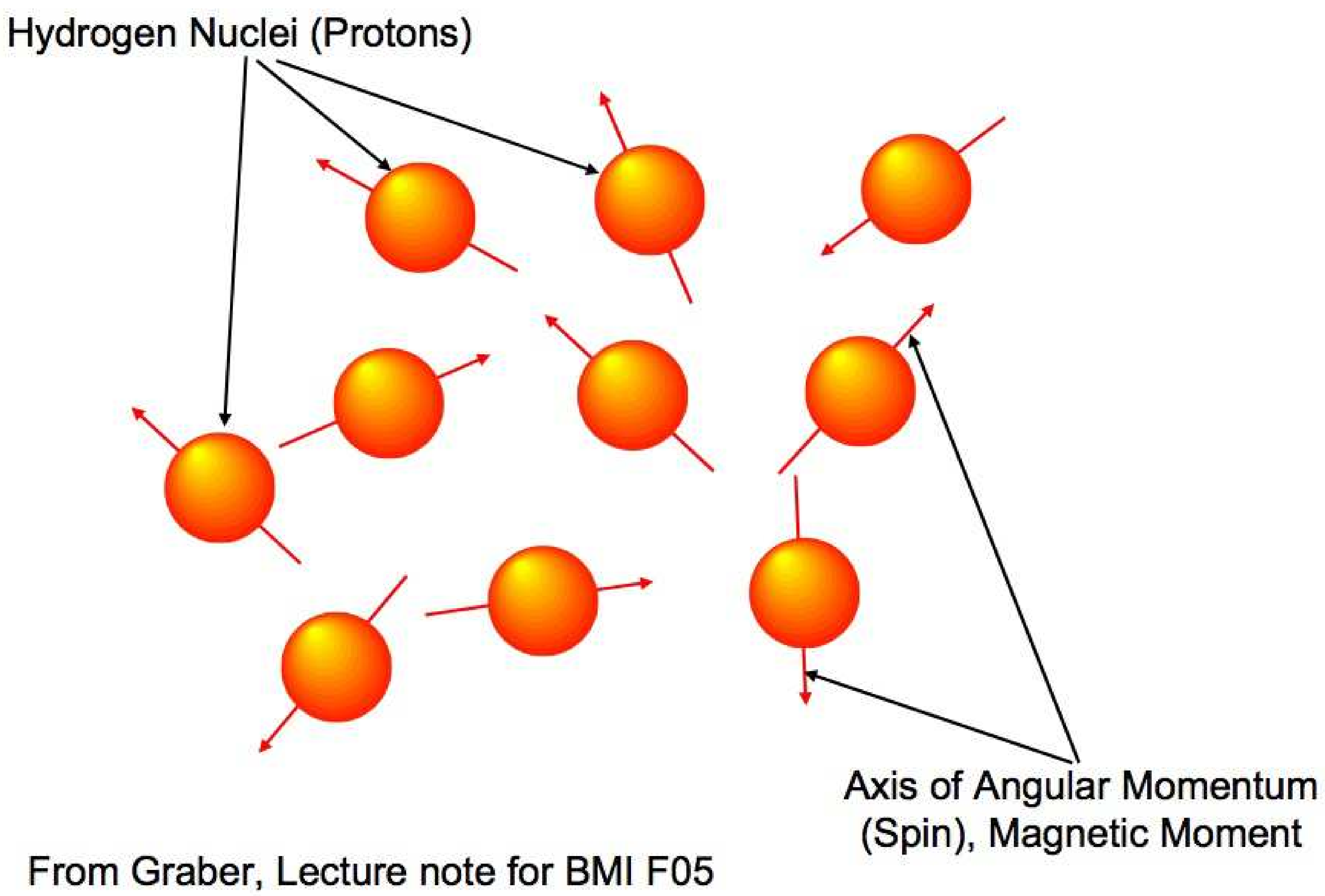

Figure 2: Naturally random spin orientation of nuclei

MR Image Acquisition

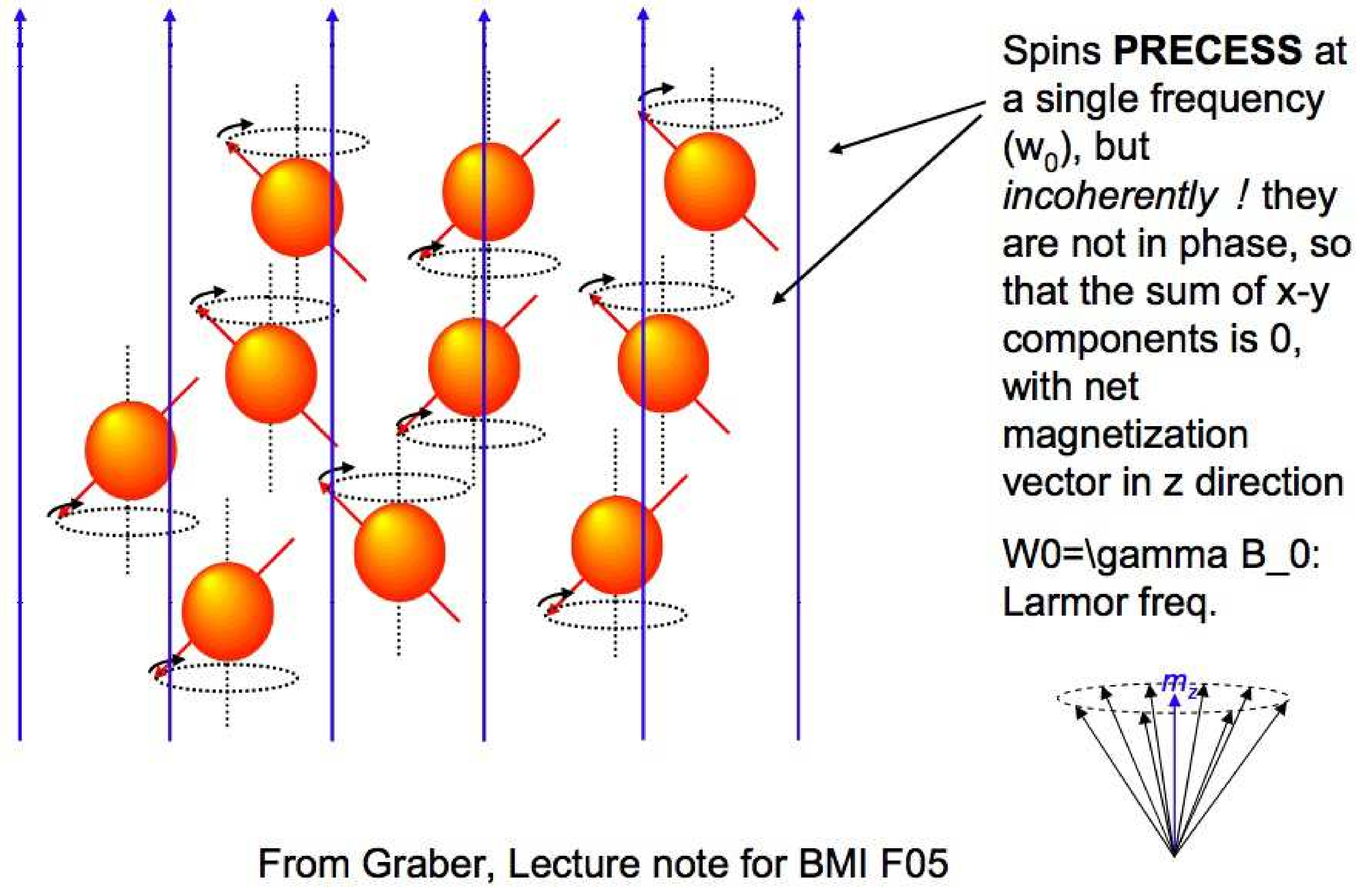

Figure 3: Spin alignment with static magnetic field

MR Image Acquisition

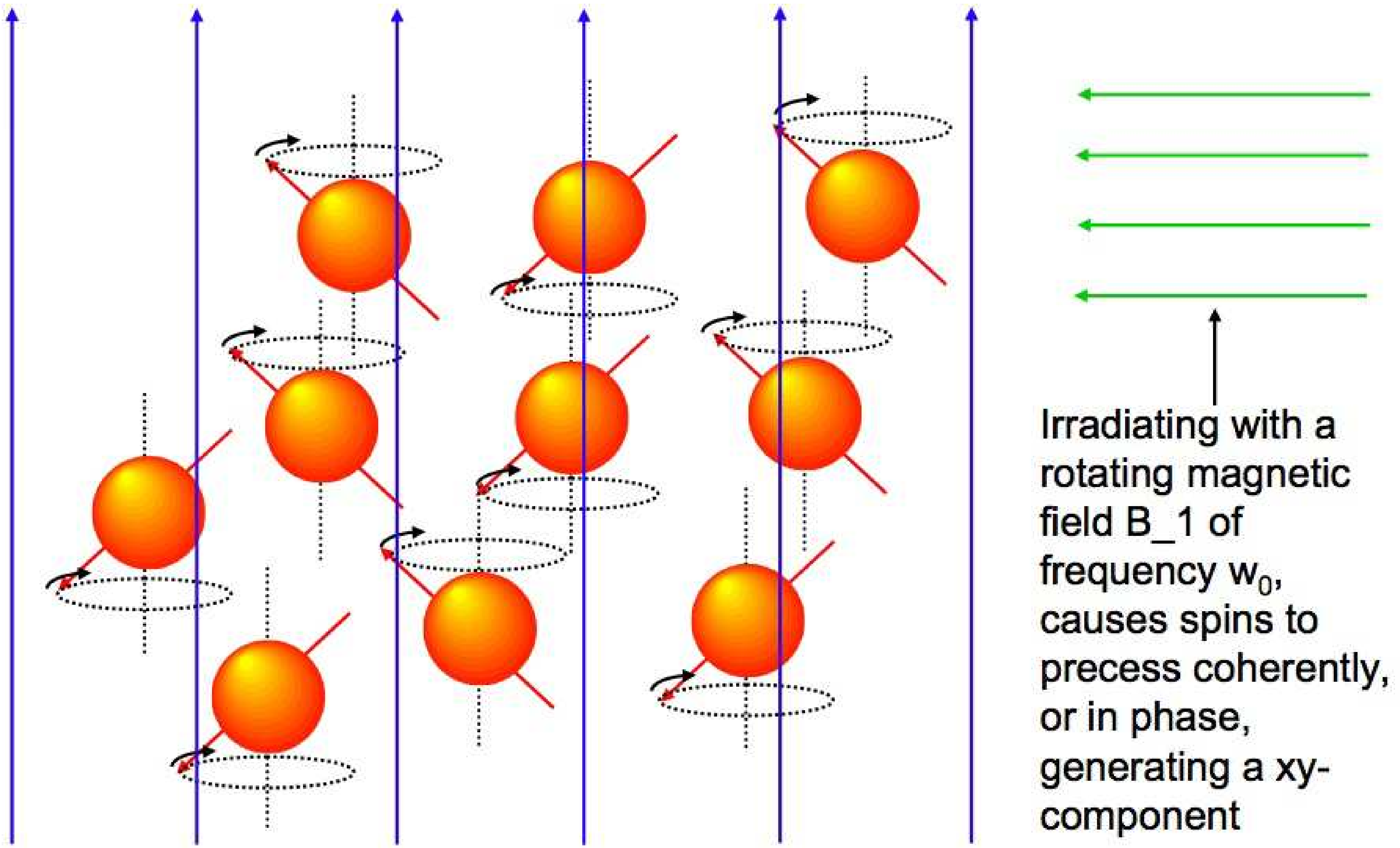

Figure 4: Excitation with rotating magnetic field

Topics

Pros and Cons of MRI

Selected Pros

- Patient care (radio frequencies from 10^6 Hz to 10^8 Hz, and wavelengths from 1 Hz to 100 Hz, no X-ray!)
- High spatial resolution (50 µm and lower)
- Excellent contrast resolution (discrimination of soft tissues)
- Rich information about anatomical structure
- Anatomical reference for functional modalities
- Enables quantitative studies
- Pre-, intraoperative guidance for intervention
- Functional imaging modality (diffusion, perfusion, flow imaging)
- Enables the usage of contrast agents

Pros and Cons of MRI

Selected Cons

- Magnetic field makes devices and especially hybrid scanners expensive (investment and maintenance).
- Inhomogeneities are caused by the radio frequency coil.
- Intensity inhomogeneities produce spatial changes in tissue statistics.
- Intensity-based segmentation fails due to inter- and intra-scan inhomogeneities.
- Inhomogeneities can change with different acquisition parameters, from patient to patient, and from slice to slice.
- It is dangerous to approach the magnet with ferro magnetic objects, see for instance here.

Topics

On the Physics of MRI

Pros and Cons of MRI

Summary

- Take Home Messages
- Further Readings

Take Home Messages

- MRI physics are based on the magnetic spin orientation of the nuclei in a strong static magnetic field. Deflecting the spins with a second magnetic field enables the readout for all axes.
- There are a lot of advantages of this technology, but also some disadvantages. Intensity inhomogeneities are a major issue which is adressed in the next units.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

Bias and Gain Fields

Online Course – Unit 21

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Inhomogeneities in MRI

Mathematical Modeling

Summary

- Take Home Messages
- Further Readings

Causes for Inhomogeneities

Definition

If there are slow and nonanatomic intensity variations present in the image of one and the same tissue class, we talk about the presence of **intensity inhomogeneity** (IIH).

In other words, in an image subject to IIH it can happen, for instance, that water molecules have different intensity values in the image domain.
As a consequence no mapping of intensities to tissue classes is possible.

Major reasons for intensity inhomogeneities in MR imaging are:

- non-uniform radio-frequency,
- inhomogeneity of the static main field,
- patient motion.

Bias and Gain Fields

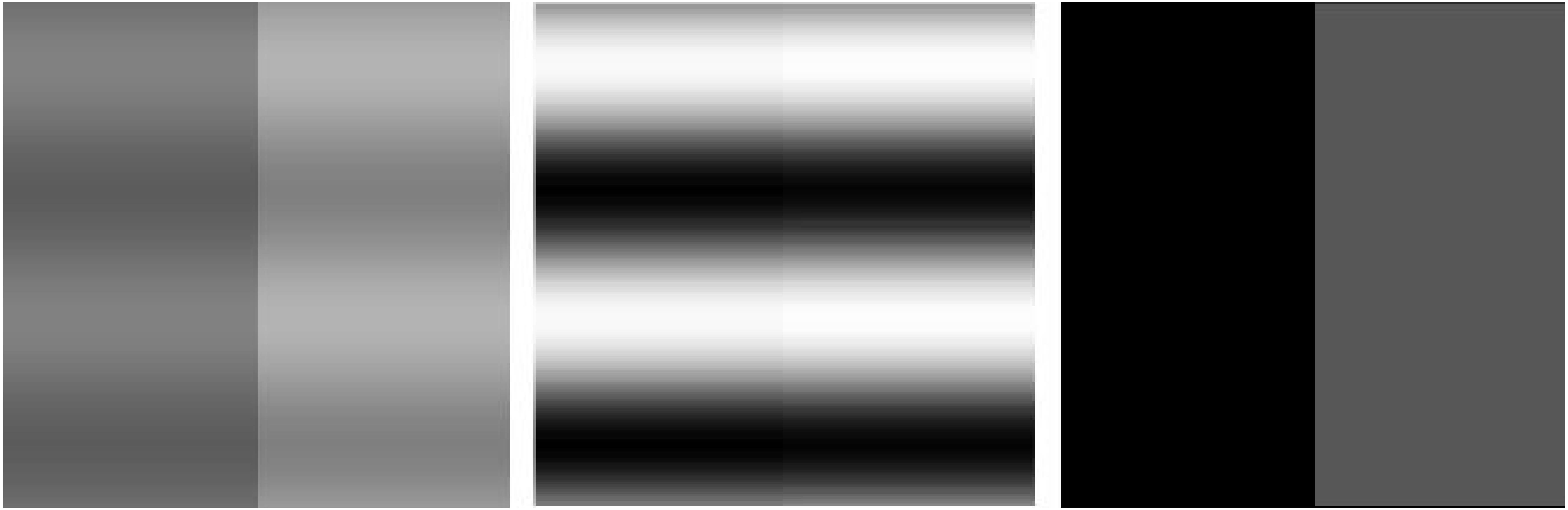

Figure 1: Observed image (left), gain field (middle) and ideal image (right) (image courtesy of W. Wells, Harvard University)

Bias and Gain Fields

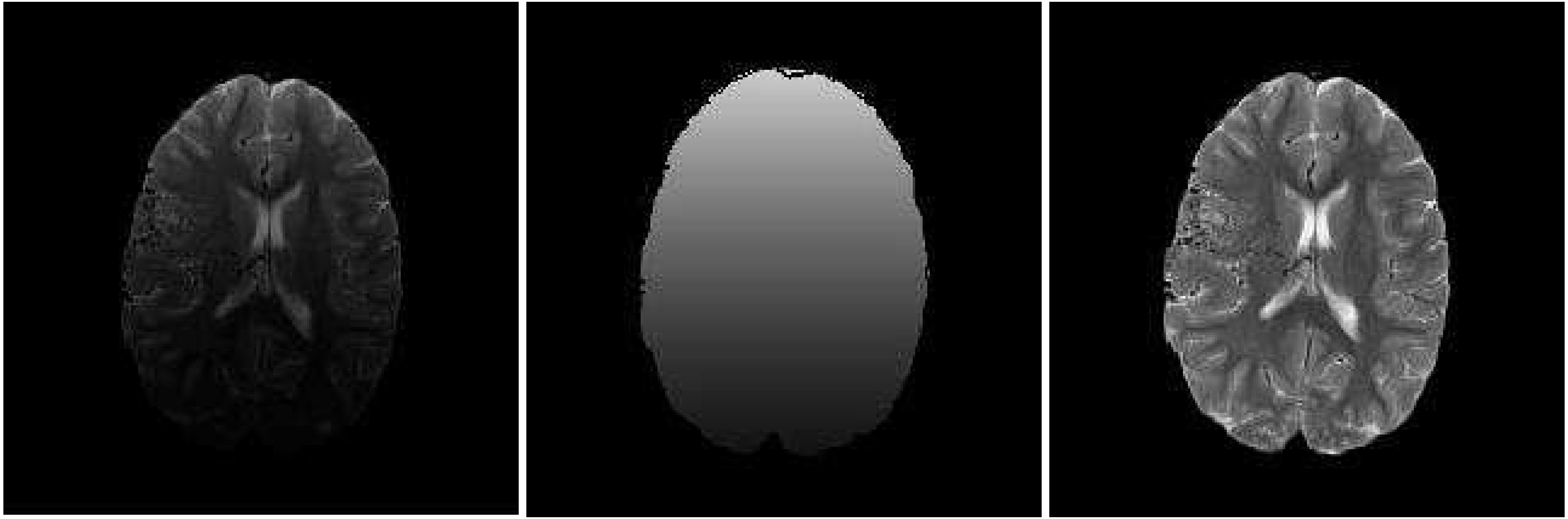

Figure 2: Original MR image (left), gain field (middle) and restored image (right)

Segmentation With and Without Intensity Correction

In this figure the middle image shows that naive segmentation using gradient information can produce inacceptable results. Only appropriate preprocessing implies a segmentation result as shown on the right.

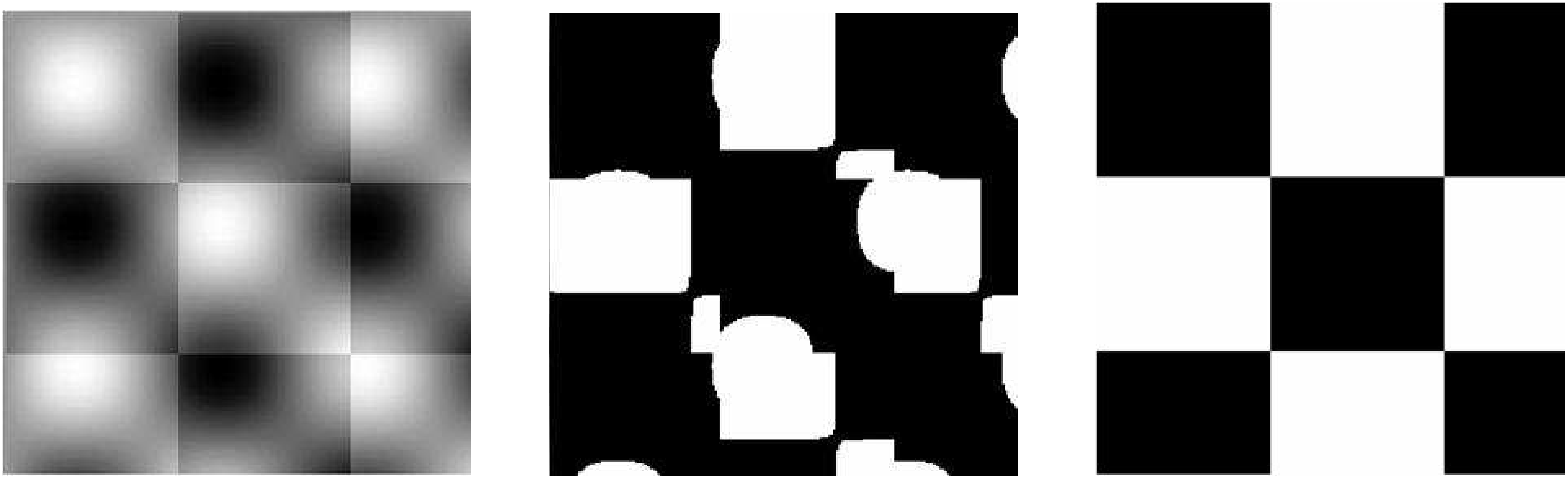

Figure 3: Biased checkerboard (left), segmentation result (middle) and ideal segmentation (right) (image courtesy of W. Wells, Harvard University)

Thresholding and Inhomogeneities

Even simple thresholding ideas fail due to intensity inhomogeneities:

Figure 4: Different MR images are binarized by the same threshold. It is obvious that heterogeneity maps identical tissue classes to different intensities. (Florian Jäger, Pattern Recognition Lab, FAU)

Topics

Inhomogeneities in MRI

Mathematical Modeling

Summary

- Take Home Messages
- Further Readings

Categories of Mathematical Models

Mathematical models for IIH are classified in three main categories:

1. **Low-frequency model:** It is assumed that IIH is caused by low-frequency components. The IIH map can be recovered using low-pass filtering.
2. **Hypersurface model:** The IIH map is represented by a smooth (low-frequency) parametric function. It can be recovered by least-square-fitting (regression).
3. **Statistical model:** The IIH map is represented by a stochastic process. Depending on the selected statistical model, the IIH map can be recovered by parametric or non-parametric statistical estimation.

Mathematical Model: Gain Field

Definition

The **gain field** $b = [b_{i,j}]$ is modeled as a multiplicative field which is applied to the ideal image $f = [f_{i,j}]$ pixelwise, i. e., all intensity values $f_{i,j}$ are multiplied by a spatially varying factor $b_{i,j}$:

$$g_{i,j} = f_{i,j} b_{i,j} + n_{i,j}, \quad \text{for } i,j = 0,1,\ldots,N-1,$$

where $g_{i,j}$ denotes the observed intensity at grid point (i,j), and $n_{i,j}$ is additive Gaussian noise.

Note: IIH correction is mostly applied to the product $b_{i,j} f_{i,j}$. Usually, Gaussian noise is eliminated before by low-pass filtering, smooth model fitting, or regularization.

Mathematical Model: Bias Field

Definition

The **bias field** ${}^{\log}b = \left[{}^{\log}b_{i,j}\right] = [\log b_{i,j}]$ results from logarithmizing the gain field. In the absence of additive Gaussian noise this results in an additive model:

$$\log g_{i,j} = \log f_{i,j} + \log b_{i,j}, \quad \text{for } i,j = 0,1,\ldots,N-1.$$

Some researchers also incorporate additive noise in the above equation with logarithms, but this is not a proper noise model for practical applications.

Mathematical Model: Bias/Gain Correction

- **Gain field correction:** If the gain field is known, the computation of the ideal image can be done pixelwise by:

$$f_{i,j} = \frac{g_{i,j}}{b_{i,j}}, \quad \text{for } i,j = 0,1,\ldots,N-1.$$

- **Bias field correction:** If the bias field is known, the computation of the ideal image can be done pixelwise by:

$$\log f_{i,j} = \log g_{i,j} - \log b_{i,j}, \quad \text{for } i,j = 0,1,\ldots,N-1.$$

Multiplicative vs. Log Additive Model

- Early models for image inhomogeneity assumed a pure additive model, but additive effects are rarely observed in MRI.
- Smooth multiplicative inhomogeneity is in accordance to the physics of MRI imaging.
- The log additive model with Gaussian noise is not modeling the MRI acquisition specific noise properly.
- Today multiplicative models are commonly accepted.

Topics

Inhomogeneities in MRI

Mathematical Modeling

Summary

- Take Home Messages
- Further Readings

Take Home Messages

- In MRI intensity inhomogeneties can be present in the acquired images and can make further processing ineffective.
- Gain and bias correction are crucial, not only, but especially in MRI.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of ‘Sandy’ Wells’ publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

Filtering in Frequency Domain

Online Course – Unit 22

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Frequency Domain Filters

Homomorphic Filtering

Summary

- Take Home Messages
- Further Readings

Frequency Domain Filters

Design a high-pass filter that eliminates the low frequency bias field.

Frequency Domain Filters

Let us consider the idea of high-pass filtering first by designing a filter in frequency domain:

- First, the observed input image $g = [g_{i,j}]$ is Fourier transformed:

$$G = \mathrm{FT}([g_{i,j}]) \,.$$

- Second, a high-pass filter is defined in the discrete frequency domain by:

$$H_{k,l} = 1 - \beta e^{-\frac{k^2+l^2}{2\sigma^2}} ,$$

  where
  - β is a scaling factor that ensures that $H_{k,l} \geq 0$ for all $k, l = 0, 1, \ldots, M-1$,
  - and σ^2 is closely related to the bandwidth of the filter-kernel.

Frequency Domain Filters

The relation between low- and high-pass filters is:

$$\begin{aligned} f &= g * h_{\mathrm{HP}} \\ &= \mathrm{FT}^{-1}(\mathrm{FT}(g * h_{\mathrm{HP}})) \\ &= \mathrm{FT}^{-1}(G \cdot H_{\mathrm{HP}}) \\ &= \mathrm{FT}^{-1}(G \cdot (1 - H_{\mathrm{LP}})) \\ &= g * (1 - h_{\mathrm{LP}}) \\ &= g - g * h_{\mathrm{LP}}. \end{aligned}$$

Frequency Domain Filters

Using the convolution theorem, high-pass filtering is simply a multiplication in the frequency domain:

$$F_{k,l} = G_{k,l} \cdot H_{k,l},$$

for all $k, l = 0, 1, \ldots, M-1$.

The final output image f is obtained by computing the inverse Fourier transform:

$$f = \mathrm{FT}^{-1}\left([F_{k,l}]\right).$$

Topics

Frequency Domain Filters

Homomorphic Filtering

Summary

- Take Home Messages
- Further Readings

Homomorphic Filtering

These filtering approaches assume that IIH is

- an artifact with low frequencies, and
- anatomic structures contribute to the high frequencies of the image.

Elimination of image inhomogeneities can be done by low-pass filtering.

Homomorphic Filtering

Subtract the low-pass filtered image and normalize the mean.

Homomorphic Filtering

Homomorphic filtering is applied to log-transformed images:

- Make a low-pass filtering of the log-transformed image
  $$[h_{i,j}] = \mathrm{LPF}\left([\log g_{i,j}]\right),$$
  where LPF denotes a low-pass filter (like averaging or a Gaussian filter).
- The IIH corrected, log-transformed image $\log f$ results from the difference:
  $$[\log f_{i,j}] = [\log g_{i,j}] - [h_{i,j}] + \mu,$$
  where μ ensures that the correction is mean preserving.

Topics

Frequency Domain Filters

Homomorphic Filtering

Summary

Take Home Messages

Further Readings

Take Home Messages

- The straightforward approach for IIH correction is low-pass filtering using the Fourier transform of the image.
- When using homomorphic filtering, a similar idea is applied on the log-transformed images including a mean preservation technique.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of ‘Sandy’ Wells’ publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

Filtering in Spatial Domain

Online Course – Unit 23
Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Homomorphic Unsharp Masking

Polynomial Surface Fitting

Summary

- Take Home Messages
- Further Readings

Homomorphic Unsharp Masking

Apply mean normalization.

Homomorphic Unsharp Masking

Homomorphic Unsharp Masking (HUM) is one of the simpler IIH correction methods, and also the most commonly used approach. HUM requires the computation of:

- the global mean value μ of the intensity distorted image,
- local mean values $\mu_{i,j}$ evaluated in a neighborhood of each pixel (i, j).

If the multiplicative model is used, the estimated intensity corrected value $f_{i,j}$ is then computed pixelwise in the following manner:

$$f_{i,j} = \frac{\mu}{\mu_{i,j}} g_{i,j}.$$

Homomorphic Unsharp Masking

A few remarks on homomorphic unsharp masking:

- This IIH-correction method relies on the assumption that the local means in an image are equal to the global mean in the absence of IIH.
- Differences between the global and local means are thus caused by the bias field only.
- This assumption only holds if the neighborhood used for computing the local mean contains a representative sample of the tissue types in the image.
- The size of the neighborhood has to be chosen carefully which is an experimental problem.

Topics

Homomorphic Unsharp Masking

Polynomial Surface Fitting

Summary

Take Home Messages

Further Readings

Polynomial Surface Fitting

Approximate the low frequency bias field in the additive model by a multivariate regression polynomial and eliminate the bias by subtraction.

Polynomial Surface Fitting

The basic idea:

1. The logarithmized image is considered as a 2-D function, where the pixel coordinates (i, j) denote the sampling points and the intensities $\log g_{i,j}$ denote the associated function values.
2. Fit a parametric, smooth surface to the logarithm of the intensity values.
3. Estimate the parameters by minimizing the sum of squared differences of the surface points and the logarithmized image intensities.
4. The resulting surface is then subtracted from the logarithmic image.

Polynomial Surface Fitting

Recalling our approach for spatial distortion correction, we define now a parametric mapping for intensity undistortion:

- we consider the image point at (i, j),
- we assume separable base functions,
- and thus require univariate base functions $b_k : \mathbb{R} \rightarrow \mathbb{R}$, $k = 0, \ldots, d$, and
- the coefficients of the polynomials b_k in i and j are denoted by $u_{k,l} \in \mathbb{R}$.

Accordingly, the polynomial that approximates the bias field is defined by:

$$g_{i,j} \approx \sum_{k=0}^{d} \sum_{l=0}^{d-k} u_{k,l} b_k(i) b_l(j).$$

Polynomial Surface Fitting

The resulting least square estimation problem is:

$$[\hat{u}_{k,l}] = \arg\min_{u_{k,l}} \sum_{i,j=0}^{N-1} \left\| g_{i,j} - \sum_{k=0}^{d} \sum_{l=0}^{d-k} u_{k,l} b_k(i) b_l(j) \right\|^2 .$$

This optimization problem can be solved by computing the SVD of the associated measurement matrix.

The final bias field estimate is:

$$b_{i,j} = \sum_{k=0}^{d} \sum_{l=0}^{d-k} \hat{u}_{k,l} b_k(i) b_l(j).$$

Exercise: Compute the measurement matrix for $b_k(i) = i^k$ and $b_l(j) = j^l$, and think about a proper scaling.

Topics

Homomorphic Unsharp Masking

Polynomial Surface Fitting

Summary

- Take Home Messages
- Further Readings

Take Home Messages

- Spatial filtering can also be used for bias field correction. In homomorphic unsharp masking the local means are adapted to the global mean.
- Similar to the undistortion algorithms from earlier units, a polynomial can be used to estimate and correct the bias field.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of ‘Sandy’ Wells’ publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

Histogram-based Bias Field Correction

Online Course – Unit 24

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern
FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Entropy and KL Divergence Minimization

Summary

Take Home Messages

Further Readings

Refresher: What are histograms?

- A ***histogram*** essentially is a frequency distribution of the pixel values in an image.
- Several values in a certain range are counted into so-called ***bins*** and then frequencies are plotted against bins.
- Normalization by the total sum makes a histogram a probability distribution.

Entropy and KL Divergence Minimization

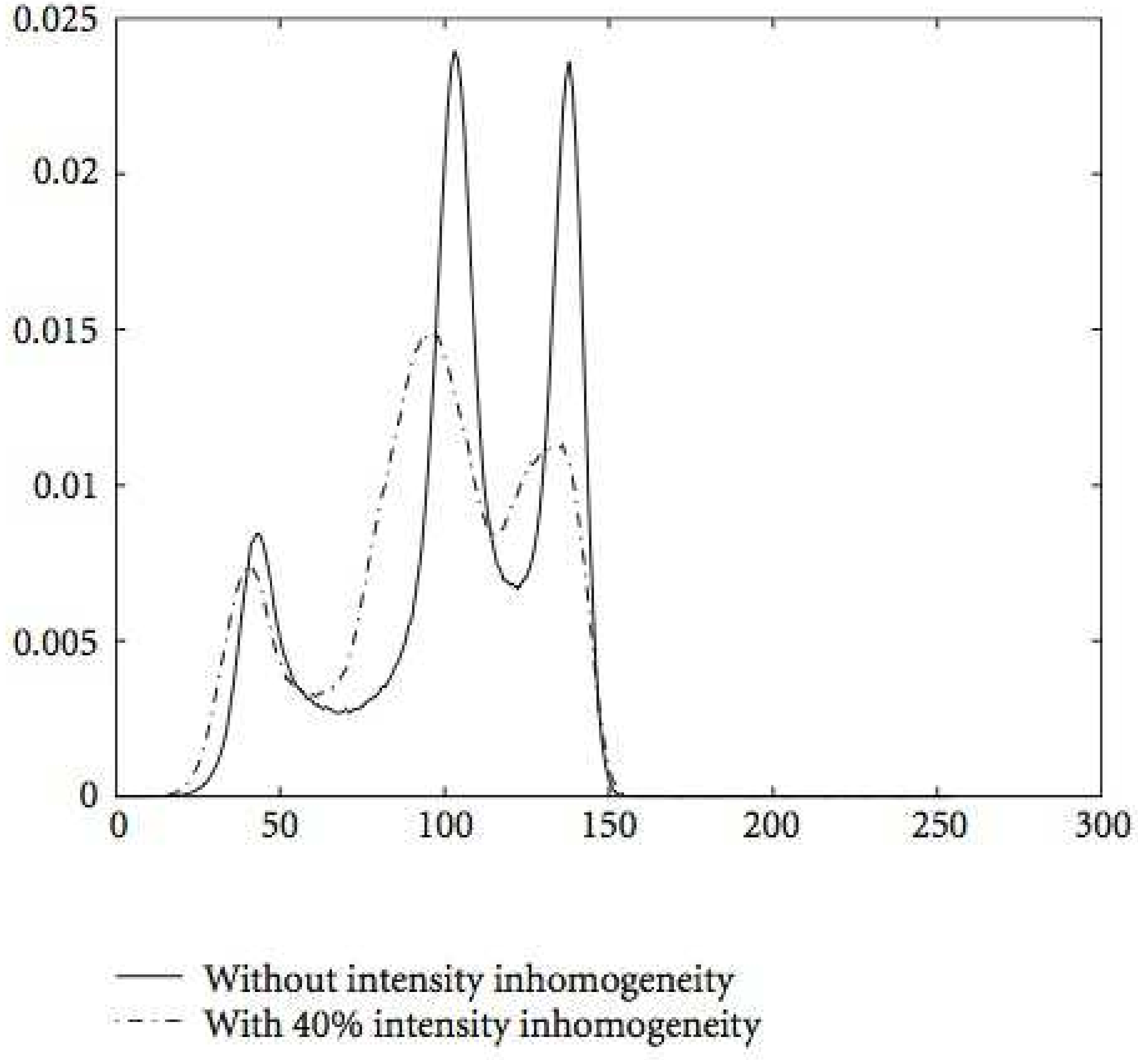

Figure 1: Histograms of brain tissue with IIH (dashed dotted line) and without bias (solid line)

Entropy and KL Divergence Minimization

Definition

The ***entropy*** H of a discrete random variable X is defined by

$$H(X) = -\sum_{i=1}^{n} p(x_i) \log p(x_i),$$

where the random measurements $x_1, x_2, \ldots, x_n$ underly the discrete probability density function

$$p(x_i) = p(X = x_i).$$

- The entropy is maximal, if the intensities are uniformly distributed.
- The entropy measures the amount of **disorder** in the image, and it is minimal for intensities with the least disorder.

Entropy and KL Divergence Minimization

Compute a bias field that minimizes the disorder in intensities.

Entropy and KL Divergence Minimization

The basic idea: Instead of fitting a parametric function with low frequencies by minimization of the squared distances between the function and the sampling values (i. e., intensities), we now apply a statistical method.

Entropy and KL Divergence Minimization

- It is assumed that the probability density function of the intensities of the original unbiased image is multimodal.
- IIH causes intensity overlap, and thus a “smearing” of the original probability density function. The peaks are no longer as sharp as they should be.
- Flattening brings the probability density function closer to a uniform density (increase of entropy).
- In terms of the statistical measures known from probability and information theory this means:
  - the bias field increases the entropy of the image,
  - the bias field decreases the Kullback-Leibler divergence between the probability density function of the image and the uniform density.

Entropy and KL Divergence Minimization

Definition

The ***Kullback-Leibler divergence*** (KL divergence) between two discrete probability density functions p and q is defined as:

$$\mathrm{KL}(p,q) = \sum_{i=1}^{n} p(x_i) \log \frac{p(x_i)}{q(x_i)}.$$

Entropy and KL Divergence Minimization

Using this definition and the functional equation of the logarithm, we find:

$$\begin{aligned}\mathrm{KL}(p,q) &= \sum_{i=1}^{n} p(x)\log\frac{p(x)}{q(x)}\\ &= \underbrace{\sum_{i=1}^{n} p(x)\log p(x)}_{H(p)=-\sum_{i=1}^{n} p(x)\log p(x)} - \underbrace{\sum_{i=1}^{n} p(x)\log q(x)}_{H(p,q)=-\sum_{i=1}^{n} p(x)\log q(x)}\\ &= H(p,q)-H(p),\end{aligned}$$

where $H(p)$ denotes entropy and $H(p,q)$ denotes cross entropy.

Entropy and KL Divergence Minimization

The KL divergence is a common similarity measure for probability density functions.

A few properties of the KL divergence:

$$
\begin{aligned}
\mathrm{KL}(p,q) &\neq \mathrm{KL}(q,p), && \text{(asymmetry)}\\
\mathrm{KL}(p,q) &\geq 0, && \text{(non-negativity)}\\
\mathrm{KL}(p,q) = 0 \quad &\Leftrightarrow \quad p = q, && \text{(definiteness)}\\
\mathrm{KL}(p,q) \to 0 \quad &\Rightarrow \quad p \to q. && \text{(continuity)}
\end{aligned}
$$

These properties show that the KL divergence is not a full metric, but resembles one, which is why it is used for the current purpose.

Entropy and KL Divergence Minimization

The algorithm for the estimation of the IIH:

- requires a parametric surface function that approximates the bias field (e. g., a bivariate polynomial), and
- the parameters are estimated such that
  - either the entropy of the resulting IIH corrected image is minimized, or
  - the Kullback-Leibler divergence w. r. t. to the uniform density is maximized.

Entropy and KL Divergence Minimization

Pros

- This approach can deal with rather complex and steep bias fields.
- It uses only a few parameters for the parametric representation of the bias field.

Cons

- Entropy minimization needs restrictions to avoid finding an all-white corrected image. This inhibits general applicability.
- The optimization task is difficult.

Entropy and KL Divergence Minimization

Note: In the Kullback-Leibler divergence based IIH correction, a **reference probability density function other than the uniform density** can be used to estimate the parameters of the bias field. In this case maximization is replaced by minimization.
This allows the incorporation of prior knowledge to IIH correction.

Reference probability density functions can be generated from:

- reference data, for instance, images of anatomic atlas data,
- ideal data that are supposed to be unbiased, for instance, manually corrected intensities,
- reference data that are acquired by costly high end MR scanners that provide superior image quality.

Topics

Entropy and KL Divergence Minimization

Summary

Take Home Messages

Further Readings

Take Home Messages

- You learned about histograms, entropy and the Kullback-Leibler divergence.
- As a rough summary, for the IIH correction entropy is minimized, KL divergence maximized.

Further Readings

The webpage of the National High Magnetic Field Laboratory can be one starting point for more detailed information regarding MRI. For an initial overview of the technology, the following article is worth reading:

MRI: A Guided Tour by Kristen Coyne.

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of ‘Sandy’ Wells’ publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. “A Review on MR Image Intensity Inhomogeneity Correction”. In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

Fuzzy C-means Clustering

Online Course – Unit 25

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

virtuelle hochschule bayern

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Fuzzy C-means Clustering

Regularization

The Regularized Optimization Problem

Summary

- Take Home Messages
- Further Readings

Fuzzy C-means Clustering

Definition

The ***fuzzy C-means objective function*** (FCM) for partitioning a set of observations into N_c classes which allows one data point to belong to more than one class:

$$J(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{N_\mathrm{c}} \sum_{k=1}^{n} a_{i,k}^d ||x_k - c_i||^2.$$

- $c_1, c_2, \ldots, c_{N_\mathrm{c}}$ are the prototypes of the clusters.
- $x_1, x_2, \ldots, x_n$ are the data points, in our particular case the logarithms of ideal intensities.
- The ***probabilistic partition matrix*** $\boldsymbol{A} = [a_{i,k}]_{k=1,\ldots,n}^{i=1,\ldots,N_\mathrm{c}}$ satisfies the probability constraint:

$$\sum_{i=1}^{N_\mathrm{c}} a_{i,k} = 1 \quad \text{for all samples } k = 1, 2, \ldots, n,$$

  where $a_{i,k} \in [0, 1]$, $i = 1, \ldots, N_\mathrm{c}$.
- The ***fuzzifier*** $d \in [1, +\infty)$ is a weighting exponent.

Fuzzy C-means Clustering

Example

Let us assume we have three clusters $i = 1, 2, 3$, and five data points $k = 1, 2, 3, 4, 5$.

For k-means clustering the partition matrix is, for instance:

$$[a_{i,k}] = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

For fuzzy C-means clustering the partition matrix is, for instance:

$$[a_{i,k}] = \begin{pmatrix} 0.5 & 0.7 & 0 & 0.1 & 0.6 \\ 0.3 & 0.2 & 1 & 0.9 & 0.3 \\ 0.2 & 0.1 & 0 & 0 & 0.1 \end{pmatrix}.$$

Fuzzy C-means Clustering

There are a few drawbacks:

- The current objective function with the probabilistic assignment of data points to classes does not consider dependencies of neighboring data points.
  Intuition: Neighboring data points most probably belong to the same class.
- The probabilistic approach requires mutually independent intensities.

The question now is, how can we incorporate dependencies of neighboring data points?

Topics

Fuzzy C-means Clustering: Regularization

Idea: Extend the FCM objective function by regularization.

- Regularization allows the
  - incorporation of *prior knowledge*, and/or
  - introduction of *penalty terms*.
- This is achieved by adding a term that biases the solution of the optimization problem towards a piecewise homogeneous labeling.

Fuzzy C-means Clustering: Regularization

A possible **regularized** objective function is:

$$J_R(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d ||x_k - c_i||^2 + \lambda \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d \sum_{x_r \in \mathcal{N}_k} \frac{||x_r - c_i||^2}{\# \mathcal{N}_k},$$

where:

- $\mathcal{N}_k$ represents the particular set of neighbors of x_k,
- $\#\mathcal{N}_k$ is the cardinality of the considered neighborhood,
- $\lambda > 0$ is a weighting factor.

Topics

Fuzzy C-means Clustering

Regularization

The Regularized Optimization Problem

Summary

- Take Home Messages
- Further Readings

The Regularized Optimization Problem

Let $[y_k]$ denote the biased logarithmized intensity values, and $[\beta_k]$ the logarithmized bias field. Now we replace the logarithm of the ideal intensity value x_k using $x_k = y_k - \beta_k$ and solve the optimization problem:

$$\left\{\hat{\boldsymbol{A}}, \hat{c}_i, \hat{\beta}_i\right\} = \underset{\boldsymbol{A}, c_i, \beta_k}{\arg\min} \left\{ \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d ||y_k - \beta_k - c_i||^2 + \lambda \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^d \sum_{y_r - \beta_r \in \mathcal{N}_k} \frac{||y_r - \beta_r - c_i||^2}{\#\mathcal{N}_k} \right\},$$

subject to the probability constraint

$$\sum_{i=1}^{N_c} a_{i,k} = 1 \quad \text{for all samples } k = 1, 2, \ldots, n.$$

The Regularized Optimization Problem

Membership evaluation:

- The optimization regarding $a_{i,k}$ has to satisfy the aforementioned probability constraint, i. e., the values must sum up to one for all k.
- For the incorporation of the probability constraint, we have to apply the Lagrange multiplier method.

Using Lagrange multipliers η_k, the extended objective function becomes:

$$J_R = \sum_{i=1}^{N_c} \sum_{k=1}^{n} a_{i,k}^{d} \left(D_{i,k} + \frac{\lambda}{\#\mathcal{N}_k} E_{i,k} \right) + \sum_{k=1}^{n} \eta_k \left(1 - \sum_{j=1}^{N_c} a_{j,k} \right),$$

where

$$D_{i,k} = ||y_k - \beta_k - c_i||^2, \qquad \text{and} \qquad E_{i,k} = \sum_{(y_r - \beta_r) \in \mathcal{N}_k} ||y_r - \beta_r - c_i||^2.$$

Estimator for Partition Matrix

The computation of the zero crossings of the gradient w. r. t. the optimized variables results in the following estimator for the partition matrix:

$$\hat{a}_{i,k} = \frac{1}{\sum_{j=1}^{N_c} \left(\frac{\#\mathcal{N}_k D_{i,k} + \lambda E_{i,k}}{\#\mathcal{N}_k D_{j,k} + \lambda E_{j,k}} \right)^{\frac{1}{d-1}}},$$

and ...

Cluster Prototype Update

... the following update for the cluster prototypes:

$$\hat{c}_i = \frac{\sum_{k=1}^{n} a_{i,k}^{d} \left((y_k - \beta_k) + \frac{\lambda}{\#\mathcal{N}_k} \sum_{y_r \in \mathcal{N}_k} (y_r - \beta_r) \right)}{(1+\lambda) \sum_{k=1}^{n} a_{i,k}^{d}},$$

and ...

Bias Field Estimator

... the following estimator of the logarithmized bias field:

$$\hat{\beta}_k = y_k - \frac{\sum_{i=1}^{N_\mathrm{c}} a_{i,k}^d c_i}{\sum_{i=1}^{N_\mathrm{c}} a_{i,k}^d}.$$

Topics

Fuzzy C-means Clustering

Regularization

The Regularized Optimization Problem

Summary

- Take Home Messages
- Further Readings

Take Home Messages

- We have seen how to use (regularized) fuzzy c-means clustering to compute an estimator of the bias field.
- The regularization is needed to incorporate local dependencies of the image data points.

Further Readings

The original paper on which the discussion in this unit is based on is:

Mohamed N. Ahmed et al. "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data". In: *IEEE Transactions on Medical Imaging* 21.3 (Mar. 2002), pp. 193–199. DOI: 10.1109/42.996338

If you want to know more about segmentation of MR images, e. g., consult

the Google Scholar record of 'Sandy' Wells' publications.

Another article worth reading is this survey paper on algorithms for intensity correction methods:

Zujun Hou. "A Review on MR Image Intensity Inhomogeneity Correction". In: *International Journal of Biomedical Imaging* 2006.Article ID 49515 (Feb. 2006), pp. 1–11. DOI: 10.1155/IJBI/2006/49515

Medical Image Processing for Diagnostic Applications

IIH Correction – Examples and Further Applications

Online Course – Unit 26

Andreas Maier, Joachim Hornegger, Markus Kowarschik, Frank Schebesch
Pattern Recognition Lab (CS 5)

FRIEDRICH-ALEXANDER UNIVERSITÄT ERLANGEN-NÜRNBERG
TECHNISCHE FAKULTÄT

Topics

Examples

Further Applications

Summary

Take Home Message

Further Readings

Bias Field Estimation: Examples

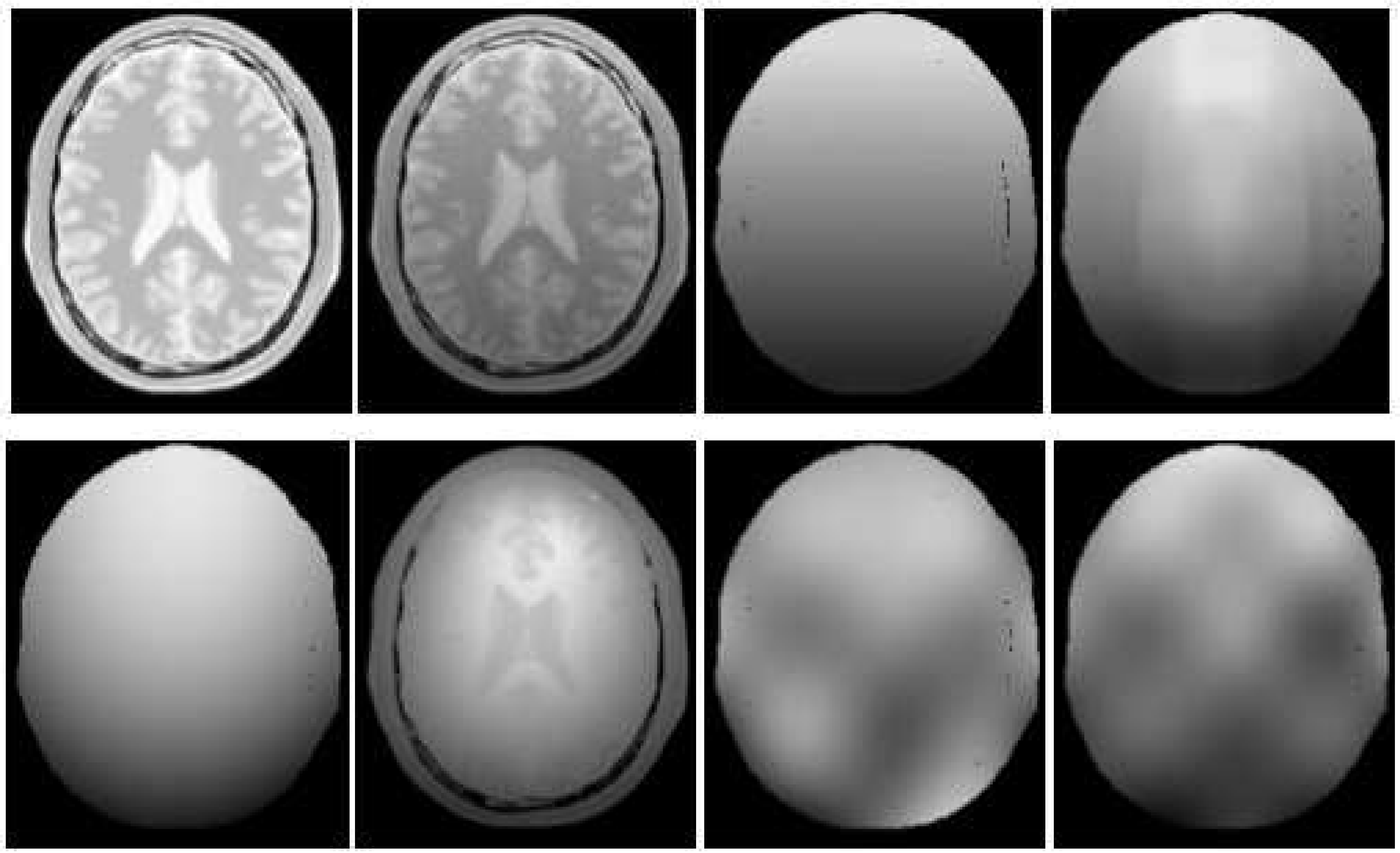

Figure 1: First row, from left to right: reference image, biased image (3% noise, 50% bias), original bias field, homomorphic unsharping mask. Second row: polynomial fit (degree 4), high pass filter, KL divergence with reference from original image, KL divergence with reference from high pass filter.

Fuzzy C-means Clustering: Image Examples

Figure 2: Comparison of segmentation results (from left to right): T1 weighted MR phantom image, fuzzy C-means algorithm without regularization, fuzzy C-means with regularization (cf. Ahmed et al.)

Topics

Examples

Further Applications

Summary

Take Home Message

Further Readings

Further Applications of IIH

- X-ray imaging: correction of the heel effect (see further readings)
- Endoscopy/retina imaging: correction of heterogeneous illumination
- Ultrasound imaging: correction of signal decay with distance from probe and correction of shadows

Further Applications of IIH

Bias correction in retina image processing:

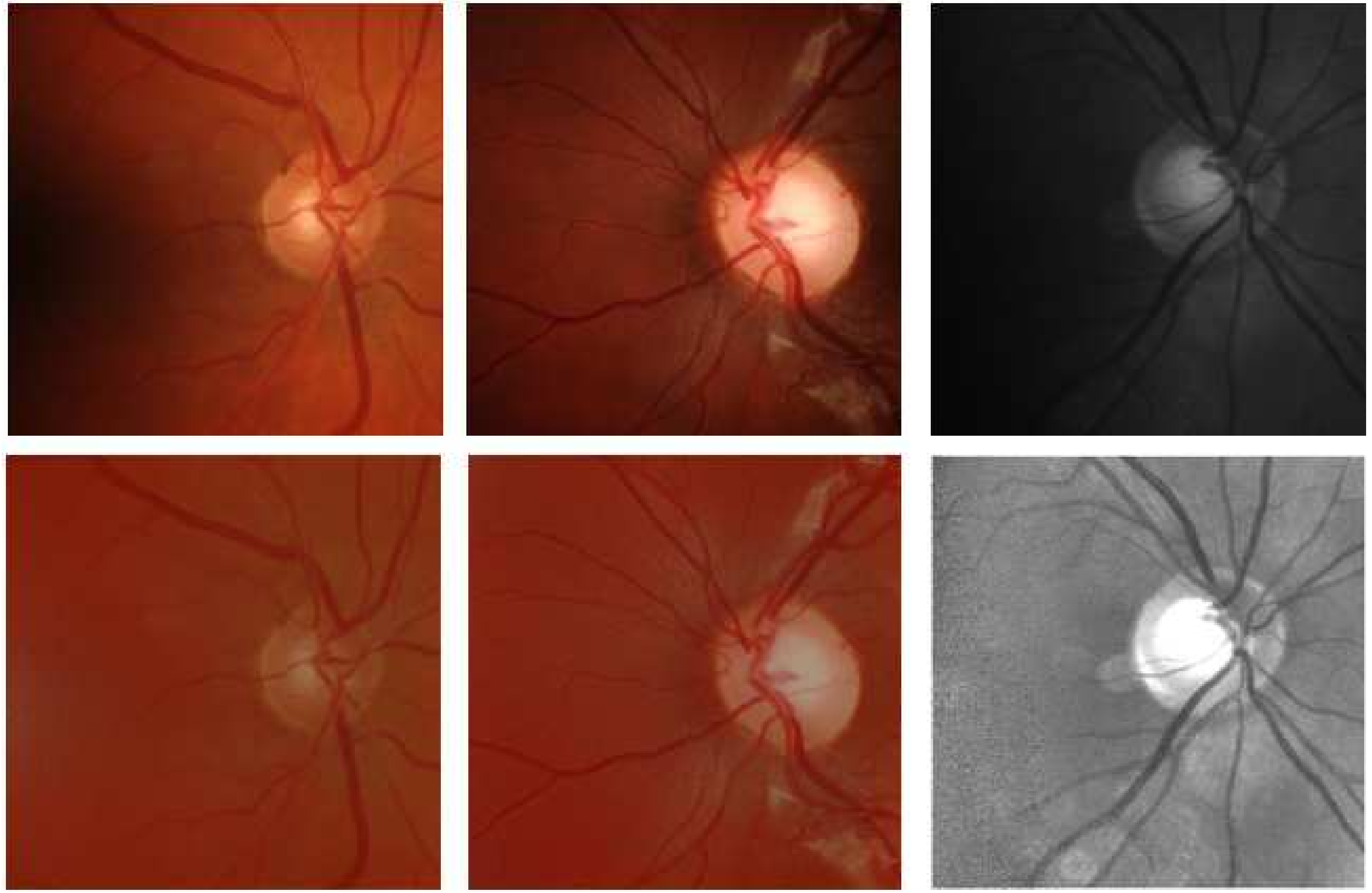

Figure 3: Retina images with heterogeneous illumination (upper row), bias corrected images (surface fitting method, degree 4 polynomials, lower row).

Further Applications of IIH

Bias correction in ultrasound imaging:

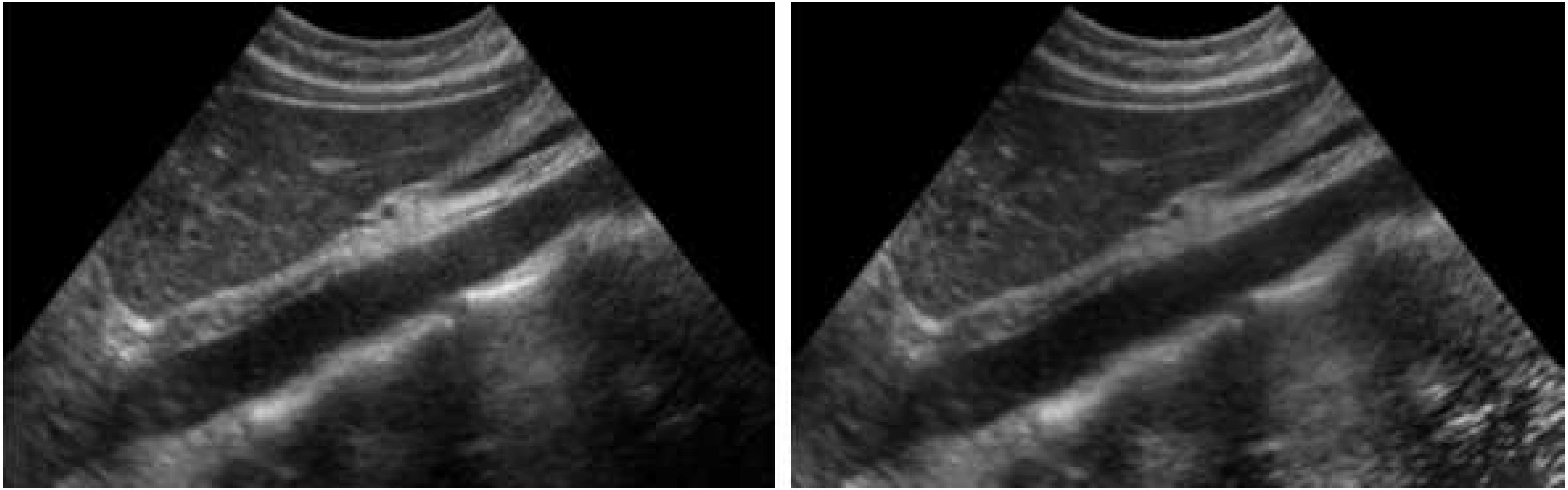

Figure 4: Ultrasound image with decreasing signal from top to bottom (left), bias corrected image (surface fitting method, degree 1 polynomial, right)

Topics

Examples

Further Applications

Summary

Take Home Message

Further Readings

Take Home Message

The correction methods introduced in this course extend their usefulness to other modalities than MRI as well.

Further Readings

The original paper on which the discussion in this unit is based on is:

Mohamed N. Ahmed et al. "A Modified Fuzzy C-Means Algorithm for Bias Field Estimation and Segmentation of MRI Data". In: *IEEE Transactions on Medical Imaging* 21.3 (Mar. 2002), pp. 193–199. DOI: 10.1109/42.996338

How the heel effect can be corrected, is described in the following paper:

Gert Behiels et al. "Retrospective Heel Effect Correction in Conventional Radiography". In: *IEEE Workshop on Mathematical Methods in Biomedical Image Analysis, 2001*. IEEE, Dec. 2001, pp. 87–94. DOI: 10.1109/MMBIA.2001.991703