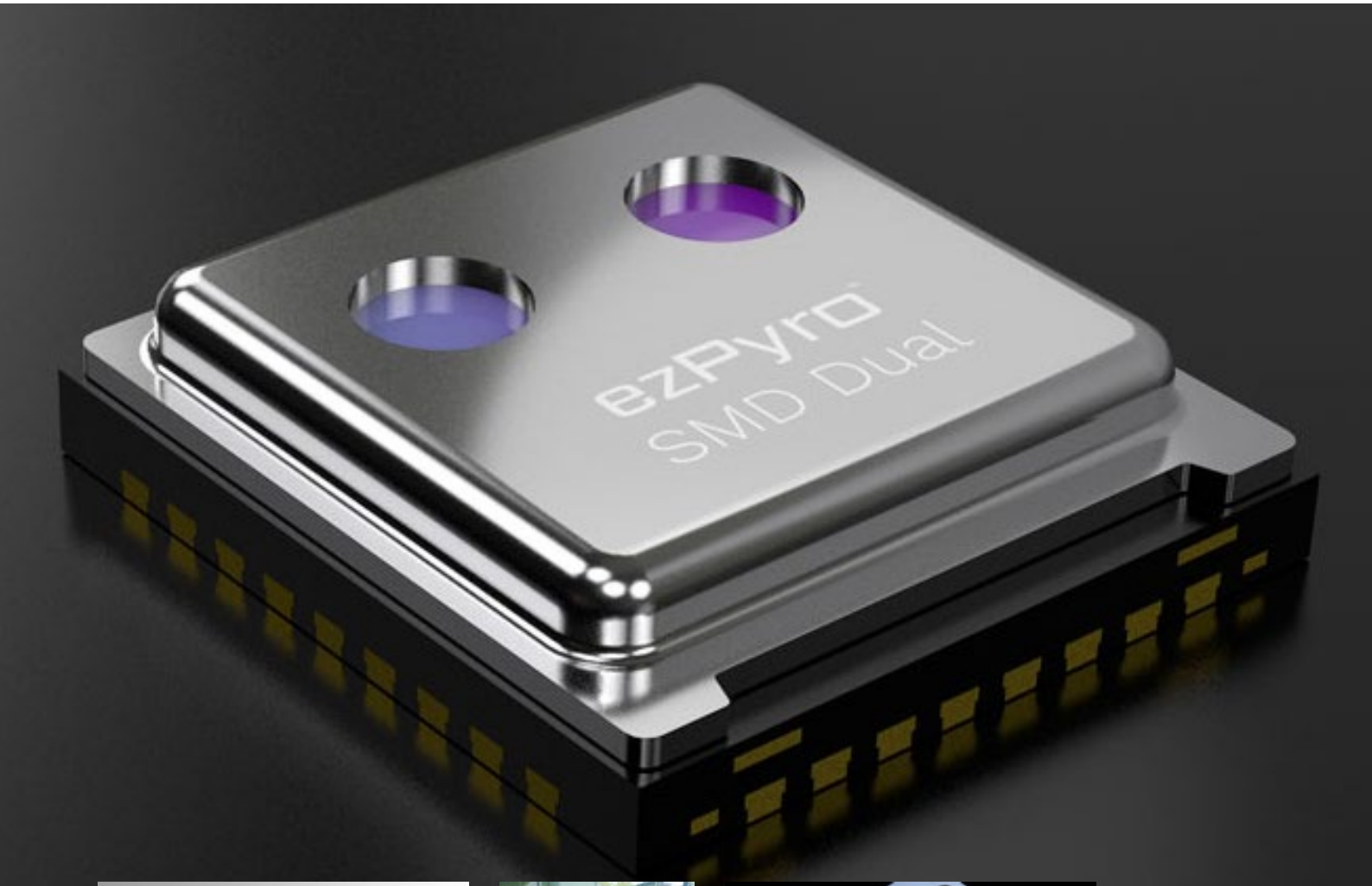
PYREOS

SENSOR INNOVATION



Guide to building mid-infrared spectral analysis system using Pyreos technology (milk example)

pyreos.com

1 Introduction

Mid-IR spectroscopy is considered special in the field of analytical chemistry as it has one major advantage over other optical spectroscopy methods and this is its capability to identify functional groups such as C=O, C-H or N-H or structural information about the sample and its composition or an identification fingerprint spectrum of a material. This can be used for solids, liquids, pastes and even gases, however it is important to note that IR spectroscopy can only be applied to molecules that contain polar bonds, i.e. molecules composed of atoms of different elements. Substances such as pure chemical elements in molecular or crystal state like O_2, N_2, Ar, Si, Ge etc cannot be measured. The spectrometer configuration depends on the sample type and conditions of the measurement. The choice of techniques such as transmission, reflection, emission and attenuated total reflection (ATR) is discussed. Key elements of the system on hardware and software levels are described.

2 ATR vs. Transmission

2.1 Transmission Technique

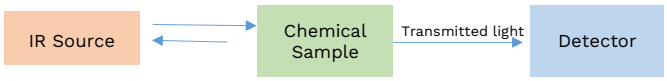


Figure: Overview of transmission spectroscopy technique

The schematic above shows a general transmission setup. Here the sample is exposed to the IR light. It can absorb some frequencies and transmit the rest. Some of the light can also be reflected back to the source. From all the frequencies it receives, the chemical sample can absorb (retain) specific frequencies and allow the rest to pass through it (transmitted light). The detector detects the transmitted frequencies, and by doing so also reveals the values of the absorbed frequencies. The output is usually measured against a background (e.g. air) and the spectrum obtained is plotted with transmission in % vs. wavelength (µm) or wavenumbers (cm^{-1}). The transmission (T) can be calculated and converted to absorbance (A) using the following equations:

$$T = (I/I_0) \times 100\%$$

Where 'I_o' is the intensity of light reaching the IR detector through the background sample (e.g. water) and 'I' is the intensity through the specimen sample (water based e.g. milk). As a ratio is calculated, instead of the intensity, the counts or volts obtained from the IR detector can also be used.

$$A = 2 - \log_{10}(T\%)$$

The sample thickness also known as the pathlength is important as the absorbance is related as shown in the equation below (Beer Lambert Law):

$$A = \varepsilon\, b\, c$$

Where ε is the wavelength-dependent absorbance coefficient, 'b' is the pathlength and 'c' is the concentration of the sample analyte.

It can be noted that absorbance is directly proportional to concentration, therefore if the setup is following Beer Lambert Law this should achieve a straight-line plot which is how the concentrations of unknown samples can be extrapolated.

The pathlength plays an important part in transmission measurements. Usually for liquids this is established using two IR transparent windows such as CaF_2 or ZnS/Se separated by a spacer of known thickness. The liquid sample can be either flowing or static in the space between the windows.

Transmission spectroscopy can also be used for solids in cases such as quality control (e.g. thickness) of plastic films where the concentration of the sample is known.

Despite its simplicity, the transmission method has some disadvantages mainly because of the challenges it entails. Firstly, it's the sample preparation, which can be complex and time-consuming, even then, there will still be inevitable accuracy and reproducibility issues due to bubbles in the flow cell when performing liquid analysis. This method is also problematic when used on samples containing water, since unlike oils, water's absorption bands in MIR are very intense, resulting in high attenuation and low signal-to-noise ratio. All the issues can be dealt with by using a different technique called Attenuated Total Reflection (ATR).

2.2 Attenuated Total Reflection (ATR) Technique

ATR spectroscopy utilises the phenomenon of total internal reflection. The basic principle is shown in the figure below.

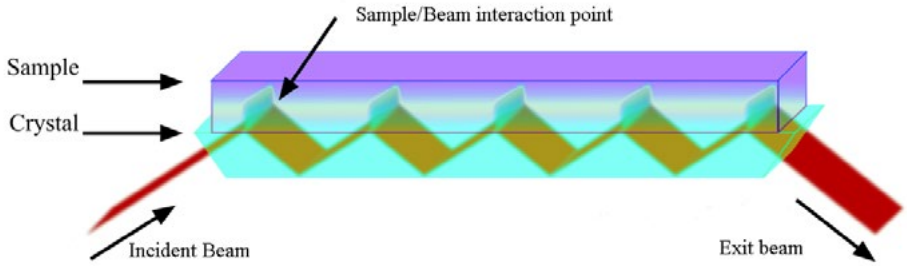


Figure: Cross-section diagram illustrating ATR spectrometer's principle of operation [1]

The light is guided through an ATR crystal, which has a high refractive index. If the incoming energy impinges on the optical surface at an angle greater than the critical angle, it will get reflected following the total internal reflection phenomenon and propagate towards the exit of the crystal. This basic effect is dependent only on the refractive indices of the optical media within and out with the crystal. An evanescent wave forms near the surface as a result of the total internal reflection. If there is an absorbing sample in contact with the crystal surface, the light forming the evanescent wave loses energy at the wavelengths where the sample absorbs. In other words, the attenuation is a result of the penetration of the electromagnetic radiation beyond the reflecting surface into the sample.

The penetration depth, defined as the distance at which the light's intensity has decreased to 1/*e* of its initial magnitude in evanescent wave, is only a few microns long, and because it depends on specific optical properties of the equipment, it can be easily controlled.

The equation for the depth of penetration (dp) is given as:

$$d_p = \frac{\lambda}{2\pi(n_1^2 \sin^2\theta - n_2^2)^{1/2}}$$

Where λ is the wavelength of the light and Θ is the angle of incidence of the IR beam relative to a perpendicular from the surface of the crystal and n_1 and n_2 are the refractive indices of the crystal and the sample being measured.

Therefore, with ATR spectroscopy the absorption's pathlength can be calculated with sufficient precision and is approximated as the penetration depth times the number of internal reflections before the light exits the crystal. Commonly, a single bounce ATR is ideal for qualitative analysis and multi-bounce crystals for quantitative analysis.

In addition, water's strong absorption is not a problem because the pathlength is not long enough for the light to be completely attenuated. The main consideration in ATR spectroscopy is that since the penetration depth is so small, very close contact between the sample and the crystal is required. In the case of water-based samples, this is not an issue as the liquid would wet the crystal, so there is no gap between them. Where solids are measured, the samples must be pressed against the crystal to ensure full contact.

Some advantages of ATR compared to transmission are:

- Minimal sample preparation. Sample does not need to undergo processing for measurement and can be measured in their natural state.
- Easy to clean after measurement
- Most suitable for thick or strongly absorbing sample such as black rubber in solids and water-based samples in liquids.
- The number of bounces in the crystal dictates the total absorbance and therefore the total accuracy of the detection.
- Can be used for both qualitative and quantitative analysis

Some disadvantages of ATR:

- Custom ATR crystal can be expensive to manufacture
- Most of the crystal materials can be scratched or harmed by corrosive substances
- Requires the sample to be homogenous
- Single bounce systems are only good for qualitative analysis
- A more complex optical setup is required
- The refractive index of the sample must be smaller than that of the ATR crystal material

2.3 Example: Measurement of Fat in Milk

When measuring fat in milk, a key challenge is sample uniformity across optical surface. Fat is like oil in water and is immiscible meaning it will float around within the liquid and may settle to the surface if left stagnant (i.e. cream).

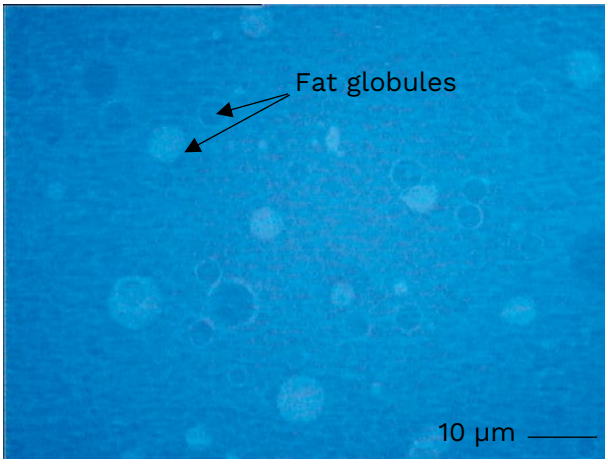


Figure: Here's an image of milk we took at 500x mag from our microscope and you can clearly see the bigger circular particles (Fat globules) and if you notice closely you will also see smaller Fat molecules around (notice them overlap - 3D). The Fat globule sizes in raw milk are usually normally distributed.

2.3.1 Measuring Milk in Transmission

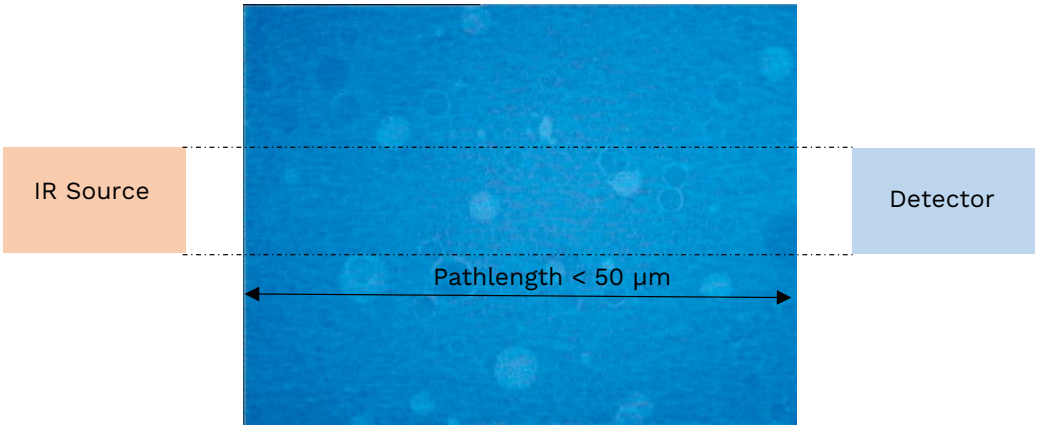


Figure: Schematic showing transmission spectroscopy example of a milk sample

Factors affecting measurement of Fat in transmission:

1. Water absorbs a lot of mid-IR which is where the Fat peak is, so the detector will receive no light
2. Need to select an optimum pathlength which is small enough for the light to pass through and be absorbed fully by water in milk. Therefore, the pathlength then becomes very critical here and needs to be precise as the absorbance is directly proportional to pathlength.
3. If the Fat globule is > 5µm light will bounce back and sample will absorb nothing (like a wall) – this is called the scattering effect and is more prominent in transmission.
4. To solve the water absorption issue, one would need a very powerful IR source (filament type) which will need to be pulsed or chopped – a chopper makes the system big, less suitable for portable applications, and prone to errors in the long run as there are more moving parts. At the same time there are no powerful pulsed IR sources to help address this issue at the moment.

Homogenised sample would be required, if not then the complete normal distribution of Fat sample should be present in the IR line of sight during a measurement scan (this is why measurement while flowing the sample can be beneficial)

Transmission cells with small pathlength < 50µm can easily get clogged and are more difficult to clean – also notice that in transmission there will be two windows that will need cleaning.

2.3.2 Measuring Milk in ATR Mode

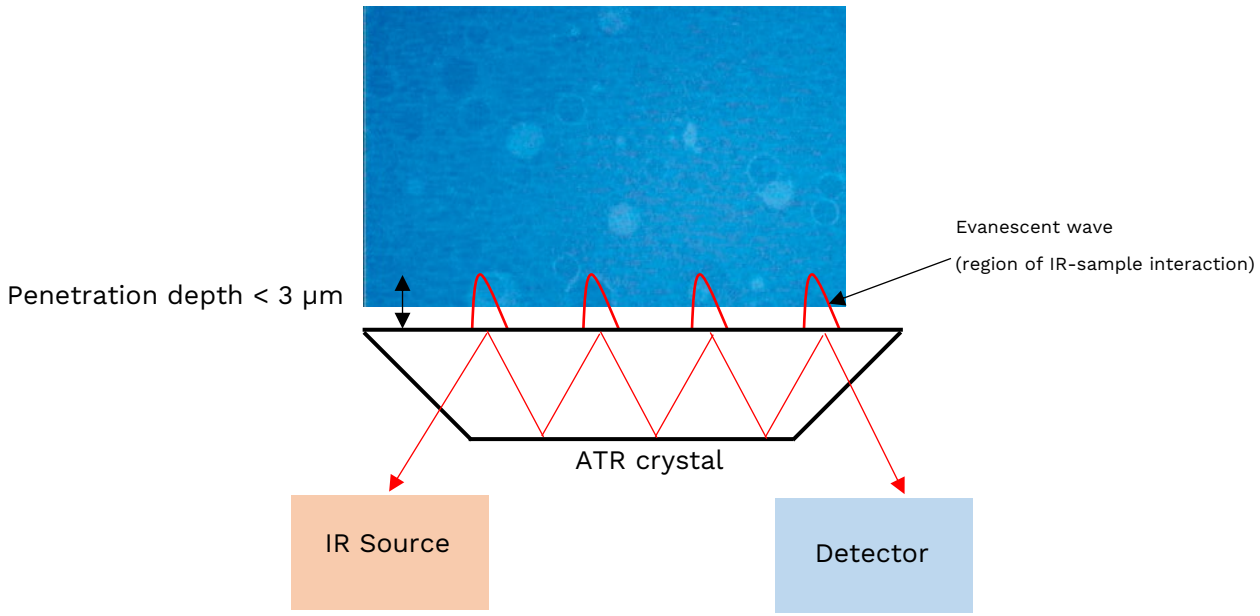


Figure: Schematic showing ATR spectroscopy example of a milk sample

In the ATR mode the only region where the IR meets the FAT globule is within the evanescent wave and if it is lucky to find the globule in that region then we would get a representative absorbance measurement from it. Conversely, if the light hits fat globules all the time (in case of a dirty crystal), the concentration of fat may always appear higher than it really is.

Summarising the following factors that would affect the Fat absorbance in ATR mode:

1. Effective penetration depth (number of bounces in the ATR crystal times penetration depth)
2. No dependence of a strict pathlength like in transmission as this is set by the refractive index of milk and crystal which does not change significantly with seasons or temperatures.
3. Homogeneity of sample (approx. sample size < 5 µm but all should be roughly the same size).
4. Less scattering effect in ATR compared to transmission
5. Do not need a very powerful pulsed IR source as compared to transmission.
6. ATR crystal will need to be very clean as any dirt on the sample/ATR interface will absorb IR.

3 Emitter Choice

Pyreos IR detectors are pyroelectric in nature, which means they possess very high sensitivity, high frequency response and do not require cooling. Pyroelectric detectors only respond to change in incident energy, therefore the IR emitter required for the system needs to have modulating IR output to observe a reasonable signal at the detector end. Usually a broadband emitter whose output follows a black body emission curve is chosen. These sources are often referred to as pulsed, 'blackbody', or thermal infrared emitters.

The following parameters when usually compared when choosing such emitters:

1. Physical size
2. Electrical input power of the emitter
3. Output IR power curve showing the maximum response at various input voltage vs. wavelengths (blackbody curve)
4. If it covers the wavelength of the application
5. Operating voltage
6. Power consumption
7. Lifetime
8. Optical Collimation

In case of Pyreos detectors there are two possible choices for emitters:

1. Steady state emitter with an optical chopper
    a. Usually filament based
    b. High thermal IR output
    c. Higher power consumption
    d. Lower lifetime
    e. Require simpler circuitry to power up emitter and chopper
    f. Involves mechanical components – a rotating optical chopper
    g. Relies on the optical collimation to enhance IR gain
2. Emitter capable of being pulsed electrically
    a. Usually semiconductor thin film (MEMS) based design
    b. Higher lifetime which depends on the pulse rate (usually 100,000 hours of operation)
    c. Lower power consumption and therefore output power (cost effective)
    d. Relies on the optical collimation to enhance IR gain (usually parabolic)
    e. Pulsed power circuitry not preferred by manufacturers (more complex than stead state operation)
    f. Can also be operated at steady state (but not preferred, as it requires other means of modulation)
    g. Allows designs with no moving parts, which may be beneficial for robustness, stability and lifetime of the solution

In case of oil or water based samples (ATR or transmission) it has been observed that the optimum emitter modulation frequency of the emitter is between 5 Hz and 10 Hz, therefore pulsed emitters are recommended for the majority of the spectroscopy setups.

Following are some emitter recommendations compatible with the 128 pixel line array detector.

Steady State Emitter Examples:

1. Boston Electronics
    a. http://www.boselec.com/product-category/ir-uv-sources/
2. Hawkeye Technologies
    a. http://www.hawkeyetechnologies.com/source-selection/#steady

Pulsed Emitters

1. Axetris
    a. https://www.axetris.com/en/infrared-sources/products
    b. Recommended: EMIRS 200
2. Hawkeye Technologies
    a. http://www.hawkeyetechnologies.com/source-selection/#pulsable
    b. Recommended: IR-75
3. Micro-Hybrid
    a. https://www.micro-hybrid.de/en/products/ir-components-and-systems/ir-sources.html
    b. Recommended: JSIR350-4

4 ATR Crystal Choices and Effects

Choosing the right ATR crystal depends on a number of factors e.g. type of sample being measured, refractive index of the sample (which must always be smaller than that of the ATR crystal so that the critical angle criteria is always met), the hardness of the crystal, desired spectral range, accepted pH range for cleaning the crystal, health hazards, and the availability/cost in manufacturing.

Here are some parameters compared:

d_p (penetration depth) has been calculated at 45 deg angle and at 1000 cm^{-1} , assuming sample refractive index of sample of 1.5.

Material	Refractive index	$d_{p\ (\mu m)}$	Water Solubility, g/100g	pH Range	Hardness (kg/mm^2)	Spectral range /µm
AMTIR	2.5	1.7	Insoluble	1-9	170	1-20
Diamond	2.4	2	Insoluble	1-14	9000	0.55-20
Germanium (Ge)	4	0.66	Insoluble	1-14	800	2-17
KRS-5	2.37	2.13	0.05	5-8	40	0.65-32
Silicon (Si)	3.4	0.85	Insoluble	1-12	1150	1.2-9
Zinc Sulfide (ZnS)	2.2	3.86	Insoluble	5-9	240	0.37-14
Zinc Selenide (ZnSe)	2.4	2	Insoluble	5-9	120	0.55-20

The following table shows some fluids, their refractive indices and also their respective critical angle calculated for ZnSe, ZnS, AMTIR and Ge.

		Critical angle (deg)			
Liquids	**Refractive Index**	**ZnSe**	**ZnS**	**AMTIR**	**Ge**
Chlorodifluoromethane R-22	1.26	31.7	34.9	30.3	18.4
Dichlorodifluoromethane R-12	1.29	32.5	35.9	31.1	18.8
Alcohol, methyl (methanol)	1.33	33.7	37.2	32.1	19.4
Water	1.333	33.7	37.3	32.2	19.5
Propane	1.34	33.9	37.5	32.4	19.6
Ether	1.35	34.2	37.9	32.7	19.7
Milk	1.35	34.2	37.9	32.7	19.7
Acetone	1.36	34.5	38.2	33.0	19.9
Alcohol, ethyl (ethanol)	1.36	34.5	38.2	33.0	19.9
Propylene	1.36	34.5	38.2	33.0	19.9
Acetic Acid	1.37	34.8	38.5	33.2	20.0
Trichlorofluoromethane R-11	1.37	34.8	38.5	33.2	20.0
Hexane	1.37	34.8	38.5	33.2	20.0
Alcohol, propyl	1.38	35.1	38.8	33.5	20.2
Heptane	1.38	35.1	38.8	33.5	20.2
Octane	1.4	35.7	39.5	34.1	20.5
Paraldehyde	1.405	35.8	39.7	34.2	20.6
Decane	1.41	36.0	39.9	34.3	20.6
Dodecane	1.41	36.0	39.9	34.3	20.6
Ethylene glycol	1.43	36.6	40.5	34.9	20.9
Propylene glycol	1.43	36.6	40.5	34.9	20.9
Chloroform	1.44	36.9	40.9	35.2	21.1
Carbon tetrachloride	1.46	37.5	41.6	35.7	21.4
Oil, olive	1.46	37.5	41.6	35.7	21.4
Furan	1.47	37.8	41.9	36.0	21.6
Glyserine (Glycerol)	1.47	37.8	41.9	36.0	21.6
Oil, vegetable 50°C	1.47	37.8	41.9	36.0	21.6
Oil, turpentine	1.47	37.8	41.9	36.0	21.6
Turpentine (wood)	1.47	37.8	41.9	36.0	21.6
Parafin, liquid	1.48	38.1	42.3	36.3	21.7
Toluene	1.497	38.6	42.9	36.8	22.0
Benzene	1.501	38.7	43.0	36.9	22.0
Oil, cedar	1.516	39.2	43.6	37.3	22.3
Ethyl salicylate	1.523	39.4	43.8	37.5	22.4
Chlorobenzene	1.525	39.5	43.9	37.6	22.4
Methyl salicylate	1.538	39.9	44.4	38.0	22.6
Ethyl cinnamate	1.559	40.5	45.1	38.6	22.9
Benzyl benzoate	1.568	40.8	45.5	38.8	23.1
Aniline	1.586	41.4	46.1	39.4	23.4
Quinoline	1.627	42.7	47.7	40.6	24.0
Carbon disulfide	1.63	42.8	47.8	40.7	24.0
Methylene iodine	1.737	46.4	52.1	44.0	25.7

It must be noted that the IR beam inside the ATR crystal must be greater than the critical angle to satisfy Total Internal Reflection. The choice of crystal material can therefore have a drastic effect on the wave propagation and hence the penetration depth inside the sample being investigated.

It can be summarised that a higher angle of incidence results in less reflections, and decreased penetration depth, lowering the overall absorbance of the spectrum. This is useful when highly absorbing or high refractive index samples are being measured.

5 Transmission Window Choices

Windows for transmission cells are usually chosen to be ZnSe or CaF_2 depending on measurement range and sample cleaning as ZnSe can be easily harmed with acids and alkaline solutions while CaF_2 is more robust but suffers from IR drop-off around 10 µm.

6 Challenges

6.1 Homogenisation

As discussed in ATR and transmission measurement mode sections, sample uniformity can play a crucial role in repeatability of the measurement. Homogenization is a mechanical treatment of the fat globules in milk brought about by passing milk under high pressure through a tiny orifice, which results in a decrease in the average diameter and an increase in number and surface area, of the fat globules. The net result, from a practical view, is a much reduced tendency for creaming of fat globules. Three factors contribute to this enhanced stability of homogenized milk: a decrease in the mean diameter of the fat globules (a factor in Stokes Law), a decrease in the size distribution of the fat globules (causing the speed of rise to be similar for the majority of globules such that they don't tend to cluster during creaming), and an increase in density of the globules (bringing them closer to the continuous phase) owing to the adsorption of a protein membrane.

6.2 Problem with Cleaning ZnSe/ZnS ATR Crystals

ZnS and ZnSe can be easily corroded with acids (releasing toxic hydrogen selenide) and alkaline solutions are known to etch its surface in both cases deteriorating its optical properties. It is therefore recommended to use detergent based pH neutral solutions or solvents such as methanol, IPA or acetone. Please note that these solvents may not be compatible with plastic pipes in the equipment.

Degreasers are designed to break up grease for easy removal. However, they are often more abrasive than a typical all-purpose cleaner. Therefore, enzymes such as Endozime AW Plus 5% have been tested and it was concluded that heating the cleaning solution up to 50°C was the most effective in removing oils and organic bio contaminants from both the ATR and transmission cells.

7 Algorithms

7.1 Choice of Algorithms

The aim of this section is to understand the limitations of spectral analysis for milk adulterant detection using a Pyreos 128-pixel line array sensor. For each option there could be various methods of chemometric analysis used and these are discussed here.

Assumptions

- All chemometric algorithms available in Matlab/Python/R are also available in other languages such as C/C++ to be compiled on embedded systems.
- Flash memory and storage capacity of the hardware running these algorithms has not been considered
- Algorithms are for adulterants in milk only

7.2 Chemometric Methods Available

These depend on whether quantification or detection/classification is used. A combination of the two methods can also be applied (i.e. detect the adulterant first and then predict its concentration (quantification)). At this moment, unfortunately there is no one method that will cover both detection and quantification.

- Adulterant classification: Predict which adulterant is in milk
- Adulterant quantification: Predict the concentration of the detected adulterant in milk

7.2.1 Methods for Adulterant Detection/Classification

Method Name	Acronym	Pros	Cons
K-Nearest neighbour	KNN	– Simple to implement – Multiclass handling – Flexible to fine tune – Only needs a few training samples per class – Works better for small number of classes	– Storage of data – Unable to detect unknown adulterant – Other methods known to perform better e.g. SVMDA – Hardware intensive
Partial least squares – discriminant analysis	PLSDA	– Uses PLS / PCA – Easy to set up – Perform better than simply PCA clustering	– Needs more samples for training – Other methods like SIMCA, SVMDA perform better
Soft Independent Modelling of/by Class Analogy	SIMCA	– Based on PCA – Easy to setup	– Separate class for each adulterant
Support Vector Machine Discriminant Analysis	SVMD	– Known to perform better than KNN – Robust – Perform better with datasets with many attributes – Less samples needed for training compared to PLSDA and ANN – Successfully used in pattern recognition applications	– Can be hardware intensive (speed and size) but less than KNN and ANN – Selection of parameters for tuning
Artificial Neural Network	ANN	– Works best if trained with lots of samples as compared to KNN, SVMDA – Works better when samples have more complex classes / peaks	– Black box type method – hard to troubleshoot – Most Hardware intensive – Performs better with more teaching samples

7.2.2 Methods for Adulterant Quantification

Method Name	Acronym	Pros	Cons
Partial Least Squares	PLS	– Ease of training – Performs best if less features/ components in sample set	– Does not perform well when other components change in system (e.g. poor adulterant detection when Fat is changing) – Relies on principal components and can be complicated to setup – Needs more samples compared to SVM for training
Support Vector Machine	SVM	– More robust compared to PLS – Works with multi parameter samples – Better performance with variations in unknown parameters (e.g. prediction of adulterant better with Fat changing)	– Tuning parameters (kernel) must be carefully selected

7.3 Conclusions on Algorithms

A combination of SVMDA or KNN along with SVM has been suggested for milk adulterant detection, classification and quantification scenarios. These methods are by no means finalised and selection depends a lot on the system. Therefore rigorous tests are suggested to validate their use for milk analysis on your device using a Pyreos line array sensor. Contact Pyreos for further support.

8 Bibliography

[1] “ATR path figure. Authored by: Fulvio314. Located at: https://en.wikipedia.org/wiki/Attenuated_total_reflectance#/media/File:ATR_path-en.svg. Project: Wikipedia. License: CC BY-SA: Attribution-ShareAlike https://creativecommons.org/licenses/by-sa/4.0/”.

PYREOS
SENSOR INNOVATION

Further resources and information at:

https://pyreos.com/resource-centre

https://pyreos.com/news

https://pyreos.com/case-studies

https://pyreos.com/sales-distribution

Contact Pyreos

https://pyreos.com/contact
+44 131 322 0732