Review

Emission Quantification via Passive Infrared Optical Gas Imaging: A Review

Ruiyuan Kang 1, Panos Liatsis 2 and Dimitrios C. Kyritsis 3,*

1 Department of Mechanical Engineering, Khalifa University, Abu Dhabi P.O. Box 127788, United Arab Emirates; 100052649@ku.ac.ae
2 Department of Electrical Engineering and Computer Science, Khalifa University, Abu Dhabi P.O. Box 127788, United Arab Emirates; panos.liatsis@ku.ac.ae
3 Research and Innovation Center on CO₂ and Hydrogen, Khalifa University, Abu Dhabi P.O. Box 127788, United Arab Emirates
* Correspondence: dimitrios.kyritsis@ku.ac.ae

Abstract: Passive infrared optical gas imaging (IOGI) is sensitive to toxic or greenhouse gases of interest, offers non-invasive remote sensing, and provides the capability for spatially resolved measurements. It has been broadly applied to emission detection, localization, and visualization; however, emission quantification is a long-standing challenge for passive IOGI. In order to facilitate the development of quantitative IOGI, in this review, we summarize theoretical findings suggesting that a single pixel value does not provide sufficient information for quantification and then we proceed to collect, organize, and summarize effective and potential methods that can support IOGI to quantify column density, concentration, and emission rate. Along the way, we highlight the potential of the strong coupling of artificial intelligence (AI) with quantitative IOGI in all aspects, which substantially enhances the feasibility, performance, and agility of quantitative IOGI, and alleviates its heavy reliance on prior context-based knowledge. Despite progress in quantitative IOGI and the shift towards low-carbon/carbon-free fuels, which reduce the complexity of quantitative IOGI application scenarios, achieving accurate, robust, convenient, and cost-effective quantitative IOGI for engineering purposes, interdisciplinary efforts are still required to bring together the evolution of imaging equipment. Advanced AI algorithms, as well as the simultaneous development of diagnostics based on relevant physics and AI algorithms for the accurate and correct extraction of quantitative information from infrared images, have thus been introduced.

Keywords: infrared optical gas imaging (IOGI); emission quantification; artificial intelligence (AI)

1. Introduction

Toxic emissions from thermal engines in the power generation sector have been the cause of serious societal concern [1], especially as they relate to greenhouse gases contributing significantly to global warming [2], which is currently the main focus of international environmental protection policies. Accurate measurement is the foundation of effective emission control and, consequently, the development of advanced, smart, and convenient emission quantification tools emerges as a pressing technical necessity [3].

Pursuing such measurements in the infrared (IR) band offers the advantage that they can directly quantify the global warming potential of gaseous emissions. Most interesting combustion species, such as hydrocarbons and species containing the C-H bond in general, including carbon oxides and nitric oxides, all have strong signals in the infrared range [4], as shown in Table 1. It is also noted that nitrogen and oxygen, which make up the majority of air, do not have IR activity because they are homonuclear diatomics, meaning there is no such interference. On the other hand, it is true that, especially in a combustion environment, IR signals are vulnerable to broad-band black body emission from high-temperature...
particles [5,6], whereas the puzzling emission spectrum of water vapor constitutes another source of interference [4,7].

Table 1. Absorption bands of several pollutants at the International Standard Atmosphere (ISA).

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Absorption Wavebands (µm)</th>
<th>Molecule</th>
<th>Absorption Wavebands (µm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>CO</td>
<td>2.29–2.48</td>
<td>NO</td>
<td>2.63–2.78</td>
</tr>
<tr>
<td>CO₂</td>
<td>2.66–2.81</td>
<td>NO₂</td>
<td>3.38–3.53</td>
</tr>
<tr>
<td>CH₄</td>
<td>3.07–3.71</td>
<td>SO₂</td>
<td>3.94–4.07</td>
</tr>
<tr>
<td></td>
<td>4.36–5.05</td>
<td></td>
<td>6.94–9.44</td>
</tr>
<tr>
<td></td>
<td>4.07–4.43</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Infrared optical gas imaging (IOGI), which involves the use of infrared imagers or infrared spectral cameras in order to generate images [4,8–10], has attracted considerable attention since the recent advances in the development of cost-effective, IR-sensitive chips [11–14]. Contrary to the well-developed field of infrared spectroscopy [15], IOGI is a multidimensional, spatially resolved measurement. This kind of optical-field measurement provides the possibility for the determination of the geometrical characteristics of pollutant dispersion. From an engineering application perspective, laser-based spectroscopy measurement has a higher cost and requires complex data interpretation by highly skilled users [16], which makes its application to industrial practice challenging. As a result, a recent comparative assessment of various optical measurements by the U.S. Environmental Protection Agency (EPA) [17] recognizes only OGI as a work practice that can be a potential alternative for the Leakage Detection and Repair (LDAR) techniques, which are currently dominant in industrial practice and are based on sampling (through “sniffers”) and ex-situ gas analysis with ionization detectors.

IOGI has now been applied in emission detection, location, and visualization. De Almeida et al. [18] reported using portable infrared cameras to detect and visualize the volatile organic compounds (VOC) leakage of four floating production, storage, and offloading facilities of a deep-offshore field in Angola. Moreover, Lyman et al. [19] used both aerial and ground-based IOGI to detect hydrocarbon emissions of more than three thousand oil and gas facilities in the Uinta Basin and reported that aerial scanning could observe the emission of 81% of gas wells, while the ground-based method could reach up to 90%. In addition, Furry et al. [20] assessed the performance of IOGI by detecting the fugitive gas leakage of six refineries which had a total of 110,000 components. In their test, they were able to scan 4500 components per hour on average and detect leakages with a minimum detection limit of 11 g per hour. They concluded that IOGI had comparable accuracy to conventional LDAR but was more efficient [21]. Because of its potential advantages in terms of high efficiency and remote sensing, IOGI has been termed a “smart” LDAR [22].

Despite these successful detection applications, the potential of utilizing IOGI for emission quantification is only starting to be explored. Some initial registries of related work have already appeared. For example, the EPA has collected information from various types of spectral cameras and recorded applications of spectral imagers for measurements of emission concentrations [23]. Additionally, Fox et al. [24] suggested the utilization of dispersion models to measure emission rates. More recently, Hagen [25] has summarized various technologies based on spectral imagers to measure column density, concentration, and emission rates. However, these reviews mainly focused on the use of spectral imagers and did not address the use of conventional IR cameras, because of reservations as to whether IR images can yield quantitative results. This constitutes a substantial shortage. Furthermore, the data processing methodologies mentioned in these reviews were limited since they are oblivious to the methods of artificial intelligence.

In this paper, we address the shortages by first analyzing the difficulties impeding IOGI emission quantification, then pointing to the fundamental fact that the value of a single pixel signal cannot provide sufficient information to retrieve the quantitative values of emission, and ultimately collecting, organizing, and analyzing the effective and potential methods for quantifying emission. In particular, we highlight the power and contribution...
of machine learning algorithms, which broaden substantially the toolbox of IOGI emission quantification and consequently improve the feasibility of such measurements.

2. Challenges in IOGI Emission Quantification

Using IR cameras to quantify emissions is recognized as a tough task. In [23], EPA claimed that the “thermal IR camera’s major drawback is its inability to measure the quantity or concentration of gas present in a gas plume”. Fox et al. [24] stated that “most current OGIs only present a qualitative (visual) flux estimate”. Hagen [25] also mentioned that “while infrared cameras have proven useful in detecting leaks, their use in quantifying leaks has only recently been analyzed, and is the subject of ongoing research”. Even the newest AI-assisted OGI system developed by FLIR, one of the top infrared camera manufacturers, can only be used for intelligent gas detection and segmentation [26].

On this condition, EPA recommends using ancillary devices for emission quantification after OGI detects and locates the emission [27]. Following this idea, Ravikumar et al. [28] used the emission factor method [29] and the Hi-Flow sampler [30] to calculate the emission rate. Almeida et al. [18] adopted an infrared gas analyzer to analyze the gas composition and concentration, Al-hilal et al. [31] utilized flame ionization detectors (FID) or photoionization detectors (PID) for gas concentration measurement, while Gal et al. [32] utilized multiple devices, i.e., infrared gas analyzer, micro-chromatography, and the accumulation chamber technique, to quantify gas concentrations and flux. In addition to utilizing sniffer and ex situ gas analyzers, Englander et al. [33,34] applied laser absorption spectroscopy (LAS) [15] to quantify column concentration. Furthermore, Lev-On et al. [22] summarized all these OGI-assisted quantification methods into five categories, i.e., the average expected, leak/no leak, random sample screening, periodic screening, and high leaker sniffing.

Figure 1 shows the simple three-layer radiative transfer modeling for IR imaging [35]. The radiation sensed by the camera is the integrated emission from three layers: background, gas cloud, and foreground. Part of the radiation is also absorbed when crossing these three layers. To simplify the analysis, scattering and reflection are ignored.

![Figure 1. Three-layer radiative transfer modeling for passive OGI.](image)

Suppose each layer is homogenous, absorbs some radiation from the prior layer, and emits radiation as well. The radiation intensity at the exit of each layer can be expressed as follows:

\[
I_{v,o,p} = I_{o,e} + I_{o,t}
\]

where \( I \) is the radiative intensity (W/(m²·sr)), the subscripts \( o, e, t \) represent the frequency of the light wave, the output of the layer, emission of the layer, and transmission of the layer, respectively.

According to the Beer–Lambert law, transmissivity can be expressed as:

\[
t_v = \exp(-k_v l)
\]
where \( t \) is the transmissivity, \( k_v \) is the absorption coefficient (\( \text{cm}^{-1} \)), and \( l \) is the light path length of the layer (cm). Meanwhile, the sum of transmissivity and absorptivity is 1 when scattering and reflection are negligible, that is:

\[
t + \alpha = 1
\]  
(3)

where \( \alpha \) is the absorptivity. Supposing all three layers are in thermal equilibrium, and following Kirchhoff’s law of thermal radiation that emissivity equals absorptivity, we substitute Equations (2) and (3) into Equation (1):

\[
I_{\nu,0} = [1 - \exp(-k_v l)]I_{\nu,B} + I_{\nu,i} \exp(-k_v l)
\]  
(4)

where \( I_{\nu,B} \) is the black body radiation, \( I_{\nu,i} \) is the radiation intensity at the incident surface of the layer. Blackbody radiation can be modeled by Planck’s law as:

\[
I_{\nu,B} = \frac{2h\nu^3}{c^2 \left[ \exp \left( \frac{h\nu}{k_B T} \right) - 1 \right]}
\]  
(5)

where \( h \) is the Plank constant, \( k_B \) is the Boltzmann constant, \( c \) is the light speed, and \( T \) is the temperature. Then, the absorption coefficient, \( k_v \), of gas species \( j \) could be calculated from:

\[
k_{v,j} = s_{v,j}(T)\phi_v(T, P)\frac{PX_j}{k_BT}.
\]  
(6)

where \( s \) is the line intensity per molecule (\( \text{cm}^{-1} / (\text{molecule-cm}^{-2}) \)), which is a function of temperature, \( \phi \) is the line shape function [36] (\( \text{cm}^{-1} \)), which is a function of both pressure and temperature, \( P \) is the local pressure (\( \text{Pa} \)), and \( X \) is a mole fraction. The total absorption coefficient can be simplified as the sum of the absorption coefficients of each species:

\[
k_v = \sum k_{v,j}.
\]  
(7)

In practice, a bandpass filter is used to select the target gas species, so Equation (7) can be simplified as follows:

\[
k_v = k_{v,t}
\]  
(8)

where subscript \( t \) represents the target gas, therefore, combining Equations (4)–(8), Equation (4) can be rewritten as:

\[
I_{\nu,0} = f(P, T, X_t l, I_{\nu,i})
\]  
(9)

The mole fraction \( X_t \) is coupled with the light path length \( l \). Using the ideal gas equation of the state, \( X_t l \), can be transformed as the product of concentration \( C_t \) and light path length \( l \), which is called column density [25,37] and can be expressed as follows:

\[
CL_t = M_t \frac{PX_t}{k_BT} l = C_t l
\]  
(10)

where \( CL_t \), \( M_t \) and \( C_t \) are the column density, molecular mass, and concentration of target gas, respectively. The final intensity at each pixel is a function of the camera characteristics and double integration of the camera incident intensity at all wavelengths of the band and the covered surface of the pixel, i.e.,

\[
I_p = f(I_{\nu,i} dv dA, D)
\]  
(11)

where \( I_p \) is the pixel value, \( f \) symbolizes a functional relationship, \( I_{\nu,i} \) is the camera incident light strength at the wavelength inside the filtered waveband, \( dA \) is a finite element of the surface that a pixel covers, and \( D \) represents the camera characteristics of transforming
radiation into pixel charge and includes chip-sensing efficiency, transforming efficiency, and noise characteristics.

In most cases, pressure can be supposed to be constant along the optical path. After accumulating the transmission and absorption of three layers, the pixel value can be represented as follows:

\[ I_p = f \left( R_b, CL_g, CL_f, T_f, T_g, D, \epsilon \right) \]  

(12)

where \( R_b \) is the background radiation, \( CL_g \) and \( CL_f \) are the column densities of the target gas substance in gas cloud and foreground, respectively, \( T_f \) and \( T_g \) are the temperatures of the gas cloud and foreground, respectively, \( D \) is the device characteristics, and \( \epsilon \) is the noise that comes from the environment and devices, such as wind effect, scattering, etc., which can be neglected and regarded as measurement uncertainty. Parameters \( R_b, CL_f, T_f, D \) and \( \epsilon \) can be summarized as environmental factor \( \epsilon_e \) as they are controlled by the environment and measurement devices and are in general considered constant for the experimental condition. Thus, for a given imager, Equation (12) can be simplified and rewritten as:

\[ I_p = f \left( CL_g, T_g, \epsilon_e \right) \]

(13)

Consequently, column density is the function of \( I_p, T_g, \epsilon_e \), that is:

\[ CL_g = f \left( I_p, T_g, \epsilon_e \right) \]

(14)

Equations (13) and (14) reveal two important insights. First, from the pixel value of an IR image, we cannot decouple the concentration and light path length—they are represented by column density. Second, the fundamental quantitative parameter, column density, is a function of three parameters, i.e., pixel value, gas temperature, and environmental factors \( (R_b, CL_f, T_f, D \) and \( \epsilon \)). Therefore, a single pixel value is not sufficient to retrieve column density.

3. Column Density Quantification

Since extracting column density from a single pixel is impossible because the pixel value is also affected by the temperature of the gas cloud and the environmental factor \( \epsilon_e \), as we have just shown, additional information is needed in order to extract quantitative information. Depending on what kind of information is added, existing column density quantification methods can be divided into two classes, i.e., elimination and augmentation.

In the elimination method, the idea is to eliminate the influence of temperature and environmental factors so that the relationship between column density and pixel value can be fixed, that is,

\[ CL_g = f \left( I_p \right); \text{ Given information about } T_g \text{ and } \epsilon_e \]

(15)

The concept of the augmentation method is to add spectral information to each pixel, that is, spectral imaging, a combination of spectroscopy and imaging, so that the image conveys temperature information as well, in a latent way. The problem then transforms to column density quantification from the spectral information at each pixel, i.e.,

\[ CL_g, T_g = f \left( SP(v), \epsilon_e \right), v \in B_g \]

(16)

where \( v \) is the frequency, \( B_g \) is the characteristic band for the target gas, and \( SP \) is the spectrum obtained at a given pixel, which is a wavenumber function.

3.1. Elimination Methods

The most straightforward idea of the elimination method is controlling all environmental factors and gas cloud temperatures, as suggested by Benson et al. [38]. Therefore, a method was proposed, whereby background temperature was controlled as constant and the gas cloud temperature was measured independently so that the recorded signal at each pixel depended on column density. This method was initially designed to calibrate the min-
imum detectable concentration of IR cameras with a known light path length. However, for emission quantification purposes, the method was heavily constrained because controlling the background characteristics proved unrealistic.

Zeng et al. [39,40] used differential pixel values instead of absolute pixel values, defined as:

\[ I_{\text{diff}} = I_g - I_e \]  

where \( I_{\text{diff}} \), \( I_g \), and \( I_e \) are the differential pixel value, pixel value of the gas cloud, and the pixel value of the environment, respectively. This differential operation is similar to the background subtraction method. With this method, the influence of environmental factors can be filtered out partially, so that \( I_{\text{diff}} \) is only dependent on the column density and the temperature difference between the environment and the plume. If we know the temperature difference, the column density can be retrieved. Compared to Benson’s method [38], knowledge of the temperature contrast is a much looser condition since there is no need to control both background and gas cloud temperatures. Therefore, Zeng’s method is more generalized and convenient to use.

Apart from using image processing to remove the influence of the background, gas-correlation imaging [41] can also achieve the same outcomes through a suitable choice of hardware settings. In short, the main principle of the gas-correlation method is image subtraction. Two emission images are generated simultaneously through one gas filter (correlation filter) and one transparent filter (reference filter). This function can be realized by either a split pupil or two separate cameras, as shown in Figure 2. The gas filter is constructed as a cylinder filled with the pure target gas that absorbs the transmission at the exact absorption waveband of the target gas in the gas cloud. Thus, the subtraction of these two images reflects the transmission or radiation of the target gas when the background temperature is much higher than the temperature of the target gas and vice versa, respectively.

![Figure 2](image)

**Figure 2.** Typical setting of the gas-correlation method: (a) using two cameras to realize gas correlation Reprinted with permission from Ref. [42], Copyright 2018, Optica Publishing Group; (b) using a split pupil to realize gas correlation Reprinted with permission from Ref. [43], Copyright 2020, Optica Publishing Group.

Based on the technology of gas-correlation imaging, several column density quantification methods have been developed. In the work of Sandsten et al. [43,44], integrated transmittance is first calculated. Similar to Zeng’s method [39], the column density can be retrieved from the integrated transmittance when the temperature contrast between the environment and the gas cloud is known. An example of this kind of relation table is shown in Figure 3. Through this method, Sandsten et al. reached a detection limit of Ammonia to 200 ppm × m under a temperature contrast between background and gas that was equal to 18 K.
Instead of using an absolute differential pixel value after gas-correlation imaging, Wu et al. [42] utilized the relative differential pixel value, which was termed normalized quantification, and defined as:

\[ I_{re} = \frac{I_{filt} - I_{ref}}{I_{ref}} \]  

(18)

where, \( I_{re}, I_{filt} \) and \( I_{ref} \) are the relative differential pixel value, pixel value in the gas-filtered image, and pixel value at the same position in the reference image, respectively. For CO quantification, it was established that \( I_{re} \) was not sensitive to temperature in the 300–400 K range at the waveband around 4.3 μm (Figure 4), so the column density was only dependent on \( I_{re} \). According to a delicate selection of filters and cameras, the detection limit of the method reached 20 ppm × m. However, this temperature insensitivity interval of 300–400 K is not suitable in several applications, including quantifying the emission of gas turbines, which may have an exhaust temperature of 800–900 K [45]. Meanwhile, the temperature insensitivity intervals of species such as CO\(_2\) and NO\(_2\) (if they exist) are different from the one of CO, thus the quantification of these species needs further exploration.

Figure 3. The corresponding relationship between integrated transmittance, temperature contrast, and column density [43]. Reprinted with permission from Ref. [43], Copyright 2020, Optica Publishing Group.

A comparison between several elimination methods in terms of their advantages and constraints is given in Table 2, in which we listed the necessary inputs, advantages, and
constraints of these methods. The molecules tested in the original studies are also listed in the table, but as universal methods, these methods can be applied to many other molecules. Though Benson's method is most straightforward, its application is highly limited, because three inputs are needed for a single column density measurement. The methods of Zeng et al. [39] and Sandsten et al. [43,44] only need two inputs by using the temperature contrast, instead of the exact temperatures of the gas cloud and the background, which makes them more practical. Wu et al. [42] removed the need for temperature input by utilizing temperature-insensitive intervals of specific species. This process decreases the amount of input, but this happens at the expense of adding more application constraints, such as being applicable to specific species, and only in narrow temperature-insensitive intervals.

Table 2. Comparison between elimination methods.

<table>
<thead>
<tr>
<th>Author</th>
<th>Molecule</th>
<th>Inputs</th>
<th>Advantages</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Benson et al. [38]</td>
<td>Methane</td>
<td>Background temperature; Gas temperature; Pixel value</td>
<td>Simple and straightforward</td>
<td>Need to control all environmental factors and the temperature of the gas cloud to create an exclusive mapping between pixel value and column density</td>
</tr>
<tr>
<td>Zeng et al. [39,40]</td>
<td>Hydrocarbon compound (e.g., ethylene or benzene)</td>
<td>temperature contrast; Differential pixel value</td>
<td>No need to control both temperatures of background and gas cloud</td>
<td>Temperature information (temperature difference) is still needed</td>
</tr>
<tr>
<td>Sandsten et al. [43,44]</td>
<td>Ammonia</td>
<td>Temperature contrast; Effective transmittance</td>
<td>No need to control both temperatures of background and gas cloud</td>
<td>Temperature information (temperature difference) is still needed</td>
</tr>
<tr>
<td>Wu et al. [42]</td>
<td>Carbon monoxide</td>
<td>Relative differential pixel value</td>
<td>Column density can be obtained without knowing temperature</td>
<td>Narrow, temperature-insensitive interval for specific species impedes the generalization of the method.</td>
</tr>
</tbody>
</table>

3.2. Augmentation Methods

As we mentioned earlier, the idea of augmentation is to add spectral information at each pixel by using spectral imaging. Quantification of species and temperature from spectra has been extensively researched [15,46,47]. A major difficulty lies in the fact that IR activity of the gases overlaps in a complicated way with black-body radiation, which makes simplified approaches, such as two-color methods unsuitable for such measurements [48,49]. There are two technologies that can handle this complicated superposition, namely, inverse modeling and, more recently, machine learning.

3.2.1. Inverse Modelling

Inverse modeling uses simulated spectra generated by radiative transfer models to approximate the experimental spectrum in order to solve for both temperature and column
density and, therefore, concentration, if the light path length is known. The procedures of inverse modeling can be divided into two steps. The first step is forward modelling, that is, building a radiative transfer model, which aims to simulate spectra as a function of both temperature and column densities of gas species, which are called spectrum control parameters $\theta$. Moreover, the influence of environmental conditions should also be modeled and included in the forward modeling process. The second step is inverse modeling by iterating or directly solving for $\theta$, based on the measured spectrum.

For forward modeling, the realization of high-fidelity synthetic spectra requires radiative transfer modeling that accounts for multiple environmental factors, such as atmospheric attenuation [50] and measuring instrument function [51]. However, in some conditions, the simple one-layer radiative transfer model which merely considers the emission of the gas cloud can obtain good agreement with the measured spectra [52]. In fact, the modeling complexity depends on the conditions of the particular application condition. Before using the forward model to approximate the measured spectrum, a background-subtraction operation is usually necessary in order to “clean” the measured spectrum. Commonly, such methods include the pixel-based method [52], where the pixel values of the background are subtracted from the emission pixels, as well as principal component analysis (PCA)-based methods [53,54]. PCA is used in order to get global background features of reduced dimensionality that can then be subtracted from the acquired images.

The inverse modeling step can be pursued with several tools, e.g., maximum likelihood estimation [55] from a statistical perspective, least-squares-based regression [54,56], and Levenberg–Marquardt optimization [57]. Table 3 provides a comparative presentation of these methods. Both maximum likelihood estimation and least-squares-based regression need sophisticated considerations in model formulation, since maximum likelihood estimation is a statistics-based algorithm, which requires the design of a model with a target metric reflecting the posterior probability $P(\theta|SP)$. The solution of maximum likelihood estimation is aimed at finding the $\theta$ which has the highest posterior probability. Least-squares-based regression needs to construct a linear relationship between $\theta$ or its variants and the spectrum. Levenberg–Marquardt, on the other hand, only needs the design of a loss function that embodies the difference between the synthetic spectrum and the measured one in order to guide the optimization process.

To some extent, the probability used in maximum likelihood estimation is also a sort of loss function, therefore, both maximum likelihood estimation and Levenberg–Marquardt optimization are iterative techniques, and, consequently, they also need the definitions of initial values, learning rate, and some regularization weights in the loss function [58]. On most practical occasions, these methods take a long time to converge, and thus, the manual parameter setting is often performed, which can have a detrimental effect on the model performance [58]. Least-square-based regression is a single-step approach and does not require a manual setting of parameters, however, it poses the challenge of transforming the nonlinear relationship between the spectrum control parameters and the spectrum itself into a linear relationship with reasonable simplifications and transformations, so that the solution can be established.

Further from estimating the average temperature or concentration along the light path, reconstructing the spatial distribution along the light path is a more ambitious objective and has been extensively discussed [58–60]. One way is to assume that the light path distribution of temperature and concentration along the light path follows an approximately known functional form, such as a Gaussian distribution of concentration and temperature as a function of the radial component in a jet flame [59]. Similar ideas have also been used in extracting quantitative information from laser-absorption spectra. As shown in Figure 5, various types of profiles, such as two-T [61,62], parabolic [63], two-peak Gaussian [64], etc., have also been used. However, the choice of a distribution function requires prior knowledge about the result of the measurement, which may be unavailable in some cases. In other cases, a priori knowledge of the functional form of the solution can be replaced by a more loose assumption, e.g., continuity of the distribution [58,60], which can be used as
a regularization term in the loss function. The solution is then found by iterating on the species distribution [65,66] in order to minimize the combined loss function.

Table 3. Comparison between different inverse modeling methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>Inverse Modeling Process</th>
<th>Advantages</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum likelihood estimation</td>
<td>Determine the spectrum control parameter vector $\theta$ (temperature and column density) that results to the largest posterior probability $P(\theta</td>
<td>SP)$</td>
<td>Statistical assessment, such as confidence.</td>
</tr>
<tr>
<td>Least-squares regression</td>
<td>Formulate a linear explicit model of spectrum control parameter vectors $\theta$, so that $\theta$ can be directly retrieved by linear fitting</td>
<td>One-step process, easy to converge</td>
<td>Because of the nonlinear intrinsic properties between spectrum control parameters and spectrum, the design of a linear explicit model may require considerable development and expertise.</td>
</tr>
<tr>
<td>Levenberg–Marquardt optimization</td>
<td>Construct a loss function to represent the difference between synthetic spectrum and experimental spectrum, minimize the loss function by updating the spectrum control parameter vectors $\theta$</td>
<td>It is straightforward enough to be applied to all kinds of models so long as a loss function can be constructed</td>
<td>Also, an iterative approach, and potentially time-consuming; initial values, weights of the various terms of the loss function, learning rate, etc. need to be set manually.</td>
</tr>
</tbody>
</table>

Figure 5. Common profiles for spatial temperature/concentration distribution along the light path: (a) Gaussian, two-peak Gaussian, and uniform with boundary layer profiles Reprinted with permission from Ref. [64]. Copyright 2019, Elsevier [64]; (b) two-T profile Reprinted with permission from Ref. [63], Copyright 2007, AIAA and Prof. R.K.Hanson [63]; (c) Parabolic profiles Reprinted with permission from Ref. [63], Copyright 2007, AIAA and Prof. R.K.Hanson [63].
3.2.2. Machine-Learning-Based Methods

Machine learning can waive the need for complex modeling, manual setting of parameters (e.g., initial values and weights of loss function), and domain knowledge. It can retrieve temperature and column density (or concentration) directly from spectral data. Ouyang et al. [67] applied Extreme Learning Machine (ELM) [68], Multilayer Perceptron (MLP) [69] to quantify all NOx concentrations from the absorption spectra of automotive emissions. Due to the high dimensionality of the data, they first used PCA or a conventional autoencoder [70] to extract features from the spectra, and then fed these features to the ELM or MLP. The best performance was achieved by the deep ELM algorithm—the estimation Root Mean Square Errors (RMSE) for N2O, NO2, and NO were 1.45, 12.72, and 24.94 ppm, respectively.

In addition to quantifying average values along the light path, Ren et al. [64] quantified the spatial distribution of species along the light path from flame radiation spectra. The method is shown in Figure 6. Synthetic spectra of mixtures of CO, CO2, and H2O were fed to MLP directly without any preprocessing and the spatial distributions of CO, CO2, and H2O were used as labels to train the model. The performance of the model was excellent, and by using the waveband of 1800–2500 cm⁻¹, the Prediction RMSEs of mole fraction of CO, CO2, and H2O were less than 0.07, 0.06, and 0.04, respectively.

![Figure 6. The schematic of Multilayer Perceptron architecture used in the work of Ren et al. Reprinted with permission from Ref. [64], Copyright 2019, Elsevier [64].](image)

Although the studies above have demonstrated excellent performance and the capability of machine learning models to produce quantitative species measurements, the exploration of diverse machine-learning tools is missing from this field. Methods such as Support Vector Regression (SVR) [71], decision trees [72], radial basis function networks, [73], and the currently popular deep learning algorithms of Convolutional Neural Networks (CNN) [74], as well as transformers [75,76] for species and temperature measurements, have not as yet found wide application. Moreover, feature engineering [77], which is an inevitable prerequisite for applying conventional machine learning algorithms, has not yet been systematically explored. As we will see below, the application of machine learning in this field is still nascent and there are major challenges and opportunities. From our review, we have come to the conclusion that the potential of machine learning for extracting quantitative species and temperature information from IR data is substantial.

3.2.3. Hardware Limitations

The hardware of spectral imaging data collection has an influence on spectral imaging quality, which consequently affects quantification performance. In general, spectral imagers can be categorized as spatial-scanning imagers, spectral-scanning imagers, and non-scanning imagers.

Spatial-scanning imagers use “whiskbroom” or “push-broom” scanning to acquire a whole spectrum from a point or a line and then generate the spectral image by scanning...
the whole field of view [78]. Different technologies have been developed for spectral-scanning devices. Filtered cameras use the intuitive method of changing filters for data collection in several spectral bands [79]; Image Multi-spectral Imaging (IMSS) realizes spectral imaging by coupling optical dispersion and moveable detectors [80]; Fourier transform infrared imaging [81] can also be classified as a spectral-scanning imager [82]. A serious issue with scanning devices is that scanning can take from several seconds to 30 min in order to capture a spectral image of a gaseous IR emission [35], which makes the application of the technique very problematic in intensely unsteady flows. In lab-scale experiments, the problem is tackled using laser excitation and performing instantaneous, spatially resolved fluorescence or Raman spectroscopy [15,83].

Non-scanning imagers, also called snapshot cameras [25], can generate spectral images without scanning. A comparison of the imaging mechanism between scanning imagers and snapshot cameras is shown in Figure 7. A spectral image has three dimensions, i.e., two spatial dimensions, x and y, and a spectral dimension, λ. So spatial scanning imagers (whiskbroom spectrometers or pushbroom spectrometers) capture a vector or a matrix of spectral dimensions with or without one spatial dimension from one imaging, so several shots (scans) are needed to cover the whole cube and generate a complete spectral image. Spectral-scanning devices capture the 2D spatial domain but need to scan spectral dimensions in order to generate a complete spectral image. A snapshot camera can capture the whole 3D space in one shot. Because the snapshot does not need scanning, it can reach video rates [84]. This speed advantage is vital for high dynamical flow analysis. Although blurs may exist in snapshot images for dynamical scenes, they are easier to process than motion artifacts [84]. The main disadvantages of snapshot cameras are complex optical architecture, heavy computations for image reconstruction, and, of course, higher cost [85].

![Figure 7](image)

**Figure 7.** The one-shot imaging ability difference between various spectral imagers Reprinted with permission from Ref. [82], Copyright 2012, SPIE and Dr. N. Hargen [82].

3.3. A Comparison of Elimination and Augmentation Methods

The overview of elimination and augmentation methods points to the following limitations of elimination:

1. The need to generate lookup tables or fitting functions from calibrating massive combinations of temperature information, pixel values, and column density, which is time-consuming.
2. The requirement for prior knowledge of temperature from ancillary devices, which raises the question of why not directly use ancillary devices instead of IOGI in order to quantify the emissions, such as all kinds of laser diagnostics.
3. Acquisition of ancillary information requires access to the emission sites, which cancels the remote sensing advantage of IOGI.
4. Since the infrared image captures black-body and rotational/vibrational emission from all molecules in the utilized bandwidth, in order to measure the column density of a specific molecule, a particular narrow-bandpass filter is necessary. The consequence is the elimination method cannot quantify several molecules simultaneously.
On the contrary, because of the previous knowledge of spectral information, the augmentation method can avoid this weakness, as shown in Ren et al. [64].

It is probably for these reasons that spectral imaging is gaining momentum as an effective tool for emission quantification. The gas cloud imager, a snapshot camera developed by Rebellion Photonics, has been funded and recognized as one of the next-generation measurement devices by the U.S. government [86]. Coupled with machine learning, it shows excellent ability in leakage detection, location, and quantification [87]. Spectral imaging has also been used to provide the quantification baseline for the Alberta Methane Field Challenge of 2019 [88]. Besides, successful applications in the quantification of gas leakage [89], flame [90], and engine exhaust [91], prove the quantification feasibility and reliability of spectral imaging. Of course, the hardware needed for augmentation methods based on spectral imagers is much more costly than IR cameras, which makes ultimately the choice between the use of elimination and augmentation approach, one that relies on a cost-benefit analysis in the context of the particular application.

It is worth noting at this stage that there are various means where ordinary IR cameras can be modified to act as spectral imagers. One intuitive way is to use a filter wheel in order to select bandpass filters in sequence [92], thus effectively converting the IR camera into a filter camera [79] with the simple mechanism of Figure 8. Fast-changing filter wheels are available on the market, but a typical changing frequency is only about 17 frames per second (fps) [93]. Unusual filter shapes based on Archimedean spirals have been developed in order to improve frequency, but the improvement is limited [94]. Similar to other spectral-scanning devices, this handcrafted spectral imager is not suitable for intensely unsteady flow.

![Figure 8. A visualization of the concept of modifying an ordinary camera into a filter camera with the installation of a filter wheel Reprinted with permission from Ref. [95], Copyright 2008, IEEE [95].](image)

Generating multiple images at the camera aperture and deploying filters in order to generate multiple images on the camera chip has also been attempted. In such a configuration, the filters are arranged in a plane and generate sub-images as shown in Figure 9 [96]. Beam splitters divide the incident light, and four optical filters generate images at different wavebands. This design can be categorized as a kind of snapshot camera [82]. The main problems are low light efficiency and low resolution; the more filters used, the more serious these problems could be, so balancing spectral resolution and spatial resolution is the core problem. In particular, for IR, identifying a sufficient number of commercial filters to split into the wavebands of interest for the measurement of specific emissions may also be difficult.

Using dispersing elements instead of filters is also a possibility. In their work, Yang et al. [97] and Olbrycht et al. [98] proposed a similar method to spatialize the spectrum into images. In [97], a linear variable filter (LVF) was designed, which contained two distributed Bragg reflectors (DBR) with a wedge between them. The DBR was composed of stacks of high and low reflective index layers. As a result, the thickness of the resonance air cavity changed continuously along one direction of the device. Consequently, the transmission wavelength varied linearly across the LVF. Thus, the image pixel along
the LVF direction contains spectral information. Figure 10 shows the structure of LVF and images generated after dispersion by LVF. Clearly, the images of C$_2$H$_2$ and CH$_4$ have totally different intensity distributions, which reflects their spectral characteristics. However, we can also see some flow patterns mixed with these spectrum signals, and it is hard to isolate between the two, which limits the application of the method to only homogenous gases.

![Figure 9](image9.png)

**Figure 9.** Schematic representation of the method of division in the aperture. A, B, C, and D denote four different mirrors. Reprinted with permission from Ref. [96], Copyright 2011, IEEE [96].

![Figure 10](image10.png)

**Figure 10.** The structure and the performance of LVF: (a) the structure of LVF; (b) The IR images of C$_2$H$_2$ and CH$_4$ generated by LVF Reprinted with permission from Ref. [97], Copyright 2019, Optica Publishing Group [97].

4. Concentration Quantification

According to the definition of column density, concentration can be decoupled from the column density estimated with the methods discussed in Section 3 once the light path length is known. In other words, coupling gas cloud geometry information with column density can provide a pixel-level concentration measurement. Those techniques where the emphasis is to acquire light-path length are categorized as geometry acquisition-based methods. On the other hand, machine learning algorithms can be used in order to estimate concentration from images captured from a designated object, such as a burner or a flare. Using machine learning, light path information can be acquired implicitly rather than explicitly as in geometry-acquisition-based methods, since the 2D information (images) of the gas cloud also conveys the 3D geometry of the gas cloud for a given object. Indeed, machine learning methods can learn inverse projection mapping, thus granting access to 3D information. Methods using machine learning algorithms to estimate the emission concentration of a given object directly from 2D images are categorized as machine-learning-based methods.

4.1. Geometry Acquisition-Based Methods

At the first level of simplicity, the geometry of the emitted plumes can be very accurately known in the vicinity of emission nozzles. Gross et al. [52] applied this idea to emission measurements from an industrial smoke flare, where the smoke-cloud size was approximated as the size of the cylinder exhaust. From the column density of the gas cloud just above the exhaust nozzle, the concentration was calculated. Although the idea was simple, the performance was acceptable: the volume fractions of CO$_2$ and SO$_2$ measured by this method were 8.6 ± 0.4% and 320 ± 23 ppmv, respectively, which were close to the in-situ measurement of 9.40 ± 0.03% and 383 ± 2 ppmv.
The precondition for the application of this method is that the geometrical information of the exhaust nozzle is known or calculated in advance. However, in many practical applications, it is hard to access this kind of data, for example, in the case of on-road vehicle emission monitoring. Obviously, the method can only estimate the concentration at the location of the emission nozzle, but it is incapable of estimating the far-field geometry of the plume, which is basically determined by the flow field; consequently, we cannot retrieve concentration far from the nozzle due to the lack of knowledge of the geometry of the plume.

A more advanced technique is 3D reconstruction. There are two kinds of technologies that have already been applied for gas cloud reconstruction, i.e., stereovision and tomographic reconstruction. In stereovision, two cameras are needed in order to formulate a stereo camera. The gas cloud is imaged from two viewpoints. Then, the data is processed in a manner that can be outlined as follows. First, a geometrical model is constructed that transforms the location of the same point between two images. Second, feature points are located and matched. Usually, the feature points are corner points that can be selected by algorithms such as the SUSAN corner detector [99] or the Harris detector [100]. Then, using a combination of correlation-based and feature-based algorithms, the feature points are matched in the images from two different viewpoints. Third, a 3D surface is reconstructed. From the disparity information at the location of feature points in the images of the two views, the 3D spatial position of the feature points can be calculated, and the surface of the gas cloud can be interpolated.

Stereo-vision has been applied to the 3D reconstruction of both emission plumes [101] and fire fronts [100,102]. The reconstruction of a fire front is shown in Figure 11. With the light-path length estimated from the reconstructed 3D geometry, the average concentration along the line of sight can be calculated if the column density is known through the methods outlined above.

![Figure 11](image-url)  
Figure 11. Surface reconstruction of the flame front in a forest fire from Reprinted with permission from Ref. [100], Copyright 2011, Elsevier [100]

In addition to stereo vision, monocular vision and multi-view have also been widely developed in computer vision research to realize 3D reconstruction [103]. In fact, using machine learning, especially deep learning, is becoming mainstream nowadays [104]. Since machine learning methods do not need complex 3D modeling, which requires substantial domain knowledge, convenient end-to-end networks, and data collection to liberate engineers while assuring good performance [104]. To some extent, stereo vision and related work is a kind of “surface” reconstruction method, since they usually suppose the object is opaque and the observed light comes from emission or reflection off the surface of the object. Although this assumption is not strictly correct for emission clouds, it does not hinder using these methods for 3D surface reconstruction, with the understanding,
of course, that information inside the gas cloud, e.g., the distribution of species, cannot be retrieved.

Tomography, on the other hand, can reconstruct the surface and internal distribution of the gas cloud. Conventional tomography methods include single-step methods such as filtered back-projection [105] and iterative methods such as algebraic reconstruction technology (ART) [106]. A detailed review of conventional tomography reconstruction methods in energy applications has been reported in [107]. As for studies using IR imaging technologies, both Donato et al. [108] and Watremez et al. [109] used infrared spectral imaging and were able to reconstruct SO₂ volume cloud [98] and 3D methane concentration distributions [99]. Tancin et al. [110] used IR Laser Absorption Imaging (LAI) in order to reconstruct the CO mole fraction in a flame. Because the flame was assumed to be axisymmetric, a single image was sufficient for reconstruction. Conventional tomography methods have the shortcoming that they may need massive reconstruction computations that involve solving large-scale super-rank-deficient inversion problems [111]. In the case of iterative algorithms, these are computation-intensive and time-consuming, especially when massive numbers of images are utilized in order to generate projections.

Machine-learning-based reconstruction algorithms can address the shortcomings of conventional methods and have been applied to tomographic reconstruction from laser absorption spectroscopy and visual images. The algorithms developed in these applications can be adapted to the IR-based tomographic reconstruction [112–119]. Published works use supervised learning algorithms, which need previously reconstructed 3D geometries as ground truth data. In these studies, summarized in Table 4, images or spectra from multiple views are used as model inputs and reconstructed 3D parameter distributions inside the gas cloud are used as the targets. The pairs of inputs and targets are fed to a machine learning model, then the model is trained by minimizing a loss function that measures the difference between the 3D structures generated by the model and targets.

Table 4. Machine-learning-based tomography reconstruction of flames.

<table>
<thead>
<tr>
<th>Author</th>
<th>Methodology</th>
<th>Input</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cai et al. [112]</td>
<td>CNN+Super SloMo</td>
<td>2D low FPS visual video</td>
<td>3D kHz visual video of flames</td>
</tr>
<tr>
<td>Huang et al. [113,114]</td>
<td>CNN</td>
<td>Visual image</td>
<td>3D flame structure</td>
</tr>
<tr>
<td>Deng et al. [115]</td>
<td>DBN RNN</td>
<td>Nonlinear Tomographic</td>
<td>Section distribution of temperature and mole concentration</td>
</tr>
<tr>
<td>Huang et al. [116,117]</td>
<td>CNN</td>
<td>Absorption spectroscopy</td>
<td>Section distribution of temperature</td>
</tr>
<tr>
<td>Huang et al. [118]</td>
<td>CNN+RNN</td>
<td>Visual image</td>
<td>Prediction of flame 3D structure</td>
</tr>
<tr>
<td>Wei et al. [119]</td>
<td>CNN</td>
<td>Pixel layers of infrared laser absorption image</td>
<td>Section distribution of Methane</td>
</tr>
</tbody>
</table>

As shown in Table 4, individual methodologies may differ. Most studies adopted machine learning algorithms based on Convolutional Neural Networks (CNN), which can extract features from high-dimensional data automatically, thus decreasing the computation cost and time. In a CNN, images from multiple views can be used directly instead of being decomposed into pixel layers as in conventional tomography methods. Moreover, methods such as Recurrent Neural Networks (RNN) [120], which process sequence data, and Deep Belief Networks (DBN) [121], which can extract features in an unsupervised way, have been used. Furthermore, the introduction of machine learning also provides extra functionalities aside from reconstruction. For instance, Cai et al. [112] added the Super SloMo model [122]
to their CNN model, which helped interpolate low-fps reconstructions into high-fps video. The interpolation performance is shown in Figure 12, in which the first row is the targets, original 3D cubes of flame, and the second row is the predictions by SloMo model. It can be observed that the interpolated reconstructions are visually identical to the targets. Huang et al. [118] added a Long-Short-Time Memory (LSTM) model, a type of recurrent network, to their CNN architecture, so that the integrated model can predict the 3D dynamics of flames. As shown in Figure 13, the first row shows the ground truth images at different times, reconstructed from the experimental images, while the second row is the predicted dynamical behavior by the model. It is shown that the predicted images are similar to the ground-truth ones.

Figure 12. The application of the SloMo model on interpolating 3D flame structure Reprinted with permission from Ref. [112], Copyright 2012, Optica Publishing Group [112].

Figure 13. The application of coupling CNN with RNN on predicting flame dynamics. Panels (a–c) show the ground-truth flames at three different times during the phenomenon, whereas panels (d–f) show the corresponding result of the model proposed in [118]. Reprinted with permission from Ref. [118], Copyright 2019, Cambridge University Press.

Machine learning methods rely on data, which implies a requirement for a substantial amount of data, i.e., input and output pairs, to produce a model capable of producing high-quality reconstructions. However, it is very difficult to acquire 3D experimental data, and thus, the most feasible way to obtain target data is by utilizing conventional methods to realize reconstruction first, which, in turn, implies a long data preparation time. With
this “data expense” in advance, once the machine learning model is tuned well, it offers substantial advantages in reconstruction speed compared to conventional methods. It should also be emphasized that since the output training data needed in the training of machine learning methods are often produced by conventional methods, the quality of the resulting machine learning models cannot exceed that of the conventional methods. Alternatively, new methods to produce training targets, or approaches that require a smaller number of training data, or even no training data, e.g., unsupervised or self-supervised learning, should be developed.

4.2. Machine-Learning-Based Methods

For an object of defined geometry, one can assume that each 2D projection of the 3D object on a particular plane is unique. Thus, it is possible for a data-driven model to learn the relationship between 2D projections and 3D structure, and subsequently, determine 3D structures from recorded projections, ultimately leading to the estimation of concentration directly from a single image. These kinds of methods are termed machine-learning-based methods.

As we mentioned earlier, merely one pixel does not provide sufficient information for emission quantification. Thus, ancillary information of temperature or spectral information is needed. In the context of this approach, the additional information is spatial distribution, i.e., the IR image. For the methods considered here, it is important that reference must be made to a particular object because different objects may generate similar patterns for substantially different concentration distributions. Also, patterns generated by a different object may exceed the domain of validity of the model trained on the given dataset.

To avoid these problems, machine-learning-based methods are usually trained and deployed on a given object, which can be represented as the equation:

$$C_o = f_o(TS_O);$$

where $C$ is the concentration, $f$ is the inference function, i.e., the machine learning model, and $TS$ is the tensor, which represents a set of images or image stacks. The subscript $o$ represents the given object.

This kind of machine-learning-based method has been applied to many image-based representations of flames, such as visual images or radical chemiluminescence images obtained with certain filters [123–130]. Although similar work has not been reported with IR images, the machine-learning-based methods and ideas developed in [123–130] can be readily extended to IOGI. Table 5 summarizes the related work, where classical machine learning models are very popular, and various algorithms have been adopted, such as Gaussian Process Regression (GPR) [131], Radial Basis Function Network (RBFN), SVR, and MLP. To assure good performance of these machine learning algorithms, feature engineering is necessary, which decreases the input dimensionality and selects the most interesting features. Feature engineering methods may include extraction of statistical features such as the values of higher-order moments, Principal Component Analysis (PCA) [132], and unsupervised learning models, e.g., DBN.

Using CNNs can waive the process of feature engineering since features are extracted automatically and embedded into the concentration estimation. For instance, Rodriguez et al. [123] used the U-Net [133], a fully convolutional network to estimate the 2D distribution of soot concentration. As shown in Figure 14, the whole network uses an encoder–decoder architecture that is full of convolution layers but without any dense layers. The whole work was in end-to-end style with the input and output being images. Usually, in such concentration quantification applications, target data (i.e., the ground truth of concentrations) are collected through chemical analysis. The concentration data obtained in this way are typically averages, rather than the spatially resolved concentration data for the whole emission cloud, i.e., scalars instead of matrices. This can then support the learning by networks with a single output so that they predict average concentration. The success of [123] in retrieving the 2D distribution of soot comes from the fact that simulated soot 2D distributions were used as labels, thus, every single visual image had an output
target matrix of the same size as the original visual image to map to. U-net, which has the same size of input and output, is appropriate for this case.

### Table 5. Machine-learning-based methods for concentration quantification.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Input Format</th>
<th>Feature Extraction</th>
<th>Methodology</th>
<th>Target</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chen et al. [124]</td>
<td>Flame visual images</td>
<td>PCA</td>
<td>GPR</td>
<td>$O_2$ concentration</td>
</tr>
<tr>
<td>Li et al. [125]</td>
<td>Flame radical images</td>
<td>Non-negative matrix factorization and texture analysis</td>
<td>Fast sparse regression</td>
<td>NOx concentration</td>
</tr>
<tr>
<td>Li et al. [126]</td>
<td>Flame radical images</td>
<td>Zernike moments</td>
<td>RBFN</td>
<td>NOx concentration</td>
</tr>
<tr>
<td>Liu et al. [127]</td>
<td>Flame visual images</td>
<td>DBN</td>
<td>SVR</td>
<td>$O_2$ concentration</td>
</tr>
<tr>
<td>Ögren et al. [128]</td>
<td>Flame visual images</td>
<td>Statistical features</td>
<td>GPR, MLP</td>
<td>$H_2$, CO, CO$_2$, and CH$_4$ concentration</td>
</tr>
<tr>
<td>Golgiyaz et al. [129]</td>
<td>Flame visual images</td>
<td>Statistical features</td>
<td>MLP</td>
<td>SO$_2$, O$_2$, NO, CO$_2$ and CO concentration</td>
</tr>
<tr>
<td>González-Espinosa et al. [130]</td>
<td>Flame radical images</td>
<td>Mean</td>
<td>MLP</td>
<td>NOx concentration</td>
</tr>
<tr>
<td>Rodriguez et al. [123]</td>
<td>Flame visual images (simulated)</td>
<td>-</td>
<td>U-Net</td>
<td>Soot 2D concentration distribution</td>
</tr>
</tbody>
</table>

**Figure 14.** The application of U-Net on mapping optical images to soot field images. Reprinted with permission from Ref. [123], Copyright 2021, Elsevier [123].

The main drawback of these machine-learning-based methods is a poor generalization. The resulting models are only applicable to a given object and they may be entirely inappropriate for other objects. To tackle this, data collection, labeling, and training need to be carried out repeatedly for application to a new object, which increases the cost of utilizing the methods. Even for an object that has been “learnt”, the model still needs to be updated since the object characteristics may change with time. For example, deterioration of a gas-turbine combustor happens all the time due to erosion, carbon deposition, or deformation. So far, there has been no work reported on adaptive prediction using these models, which makes this a promising direction for future research. Research in using
transfer learning [134], meta-learning [135], and lifelong learning [136] may be offered as solutions to tackle such generalization issues.

5. Emission Rate Quantification

In addition to concentration, the actual mass flow rate of emissions is also a technically very relevant quantity. The methods to estimate emission rate can be roughly divided into two categories: (i) propagation-speed-based method and (ii) minimum-detectable-concentration-based method. If the concentration is known via the methods that we have described in previous sections, the missing information in order to determine mass flow rate is propagation speed, which is the focus of propagation-speed-based methods. However, there are also technologies, the emphasis of which is on the use of the minimum detectable concentration in order to replace the active measurement of column density/concentration, and this parameter is further used to estimate the emission rate. These technologies are categorized as minimum-detectable-concentration-based methods here.

5.1. Propagation-Speed-Based Method

The simplest way to obtain information about the emission propagation speed is to assume emission propagation speed can be approximated by the local wind speed [137], an assumption which is weak, given the strong diffusion in the gaseous phase, but it is to some extent acceptable when the gas cloud measured is far from the point of leakage, as shown in Figure 15. Using this assumption, Watremez et al. [109] developed a method coupling 3D reconstruction with wind speed. In their work, they assumed that the emission rate was constant and considered the gas cloud between two planes (labeled as P1 and P2 in Figure 15) perpendicular to the wind speed as the control body. Therefore, the emission mass in the control body can be calculated from a known concentration and reconstructed emission cloud volume, i.e.,

$$Q = \frac{mU}{L_2 - L_1} = \frac{CVU}{L_2 - L_1}$$

(20)

where $Q$, $m$, $C$, $V$, $U$, $L_1$, and $L_2$ are the emission rate, emission mass of the control body, concentration, reconstructed volume of emission cloud, emission propagation speed (wind speed), and the distances of the two planes from the leak point shown in Figure 15, respectively.

![Figure 15](image_url)

**Figure 15.** The schematic of the emission rate calculation method based on reconstruction and wind speed. Reprinted with permission from Ref. [109], Copyright 2016, SPE [109].

An improved version of Watremez’s method was proposed by Branson et al. [138], which uses the product of column density and cross-sectional area of the emission cloud in order to estimate the emission mass of the control body, that is,

$$m = A_p \sum_{i \in CS} CL_i$$

(21)
where CL is the column density, AP is the area represented by the pixel, CS is the area of the core section of the emission cloud, and subscript i refers to the pixel index. Equation (21) assumes that the lengths Li are perpendicular to the cross-section AP. Compared to Watremez’s method [80], this method does not require the calculation of the concentration and reconstruction of a 3D geometry.

Both methods should be used for gas clouds far from the emission location (leak point) since the gas cloud propagation speed can be approximated by wind speed only when the gas cloud is well-mixed with the atmosphere. However, as we have mentioned above, this assumption is weak, which may cause huge estimation errors. Watremez et al. [109] reported that the relative estimation error of the emission rate is between 7–92%. Specifically, for the estimation of the methane flow rate of 50 g/s, the results of three measurements were 15.4, 19.3, and 3.9 g/s, respectively, which were so far away from the ground truth. However, the estimations were in modest agreement with the small emission rate of 1, 10 g/s with the maximum error of 3.4 g/s. A similar estimation error trend appeared in the work of Branson et al. [138].

A more general way of estimating the propagation speed is through optical flow algorithms [139], which calculate the apparent motion of individual pixels from consecutive frames so that the gas-cloud propagation speed can be approximated. Nagorski et al. [140] used the Horn–Schunck algorithm [139], the Lucas–Kanade algorithm [141], and the image correlation velocimetry algorithm [142] to estimate the propagation speed of flare emission. Sandsten et al. [143] utilized the correlation velocimetry algorithm and Rangel et al. [44] used the Brox algorithm [144] to estimate the leakage speed of volatile organic compounds of methane and butane.

The biggest advantage of optical flow algorithms is that they have the capability of calculating the propagation speed at any section of the emission plume and do not require the emission rate to be steady. This allows measurements of emission rate from any cross-section and not only from the far-field gas cloud. In the work of Harly et al. [145], they used optical flow to calculate the emission speed immediately above the stack exit. By coupling with the prior knowledge of the exit shape and concentration, the emission rate at the stack exit can be estimated. From their measurement, the mass flow rates of CO2 and SO2 were 13.5 ± 3.8 kg/s and 71.3 ± 19.3 g/s, respectively, which were in good agreement with in situ rates of 11.6 ± 0.1 kg/s and 67.8 ± 0.5 g/s. The optical flow method is apparently more reliable compared to the methods of [109,138].

All these methods are conventional optical flow algorithms, which follow the assumptions that pixel values are constant in sequential frames captured in a short period of time and the movement of the corresponding points/blocks is slow, in the sense that there are no multi-pixel deformations in consecutive frames. However, these two assumptions are not satisfied in some real-world scenarios, such as high-speed or reactive flow. Machine-learning-based optical flow algorithms can address these scenarios and have recently received a substantial amount of attention in image processing and computer vision. A number of both supervised [146] and unsupervised [147] algorithms have been developed, and a good overview of the subject can be found in [148,149].

5.2. Minimum-Detectable-Concentration-Based Method

The minimum detectable concentration is one of the sensitivity characteristics of an IOGI device. The existence of this sensitivity value can be utilized as a passive quantitative measurement and used for emission rate estimation. The OGI-based emission factor method [150] is a technique that is derived from the conventional leak/no leak emission factor method [29] used for estimating leakages in refineries. The emission factor is defined as the emission rate of each component in an installation. The concept of the original leak/no leak method is to set a pair of standard emission factors corresponding to leak or no-leak, respectively, for a component, such as a valve or a pump. One component leaking or not is judged from a predefined emission rate threshold, and percentages of leaked components and no-leak components can be calculated correspondingly. The
average emission factor for every component in a refinery is calculated as the average of
two emission factors, weighted by percentages of leak/no-leak components.

In the OGI-based emission factor method, the predefined leakage threshold (in g/h)
is replaced by the recorded signal corresponding to the minimum detectable concentra-
tion [150], i.e., whether the component is in leakage or not depends on whether the IR
camera can “see” it or not. Thus, threshold definition, measurement and leakage deter-
nmination are integrated into a single imaging operation. Through massive amounts of
simulations on a refinery with substantial components with various leakage levels, stan-
dard emission factors can be determined from the relationship between the percentages of
leak/no-leak components and the calculated average emission factors.

The statistical nature of the method introduces two significant shortcomings. First,
it can only be used in a region with massive amounts of components, otherwise, the
calculated emission rate may not be statistically meaningful. Second, the emission rate
is a time average value, which means that the method cannot be used in order to probe
emission dynamics.

Another minimum-detectable-concentration-based method is the OGI-based disper-
sion modeling method [151], derived from the conventional inverse-dispersion modeling
method [152]. In the original approach, the downwind measured concentration of the
emission plume and the local wind speed are used as inputs for gas dispersion modeling,
such as the Gaussian dispersion [153], and backward-Lagrangian stochastic simulation [17]
in order to infer the emission rate. In the OGI-based version, the gas cloud size outlined by
the minimum detectable concentration is used as a substitute for real concentration. Then,
this information is used in tandem with gas dispersion modeling and propagation speed in
order to retrieve the emission rate.

The main shortcoming of the method is that dispersion modeling has high uncertainty
since it is an inverse modeling process, and thus, the relative estimation error may be
up to 30%, compared to the actual emission rate [152]. To reduce the estimation error,
long-time averaging is used. Moreover, the method requires sufficient ancillary information
support, such as the structure of the emission nozzle and meteorological data, which limits
its application. In fact, both the OGI-based emission factor method and the OGI-based
dispersion modeling method have not yet been tested in real emission measurements, but
rather only in simulation studies.

Zeng et al. [40] proposed an interestingly different method, which can be regarded as
a variant of the OGI-based dispersion modeling method to some extent. In this approach,
they used the size of the gas cloud, outlined by the minimum detectable concentration,
and also pixel value levels to estimate the emission rate, while information related to the
speed of propagation and dispersion modeling was not used. One would intuitively think
that omission of the velocity term would render estimation of the emission rate impossible.
However, since these are images of specific objects taken from a fixed distance, the size
of the object and signal intensity can be related to the magnitude of the emission rate.
A further simplified version of this method only used the size of the gas cloud size outlined
by the minimum detectable concentration to classify the working mode of a truck [154],
Implied by the emission levels. There are similarities to the machine learning methods
mentioned above in that the area and pixel value levels can be regarded as features selected
from the images to estimate the emission levels of a given object. However, In Zeng’s
method [40], the maximum error of the prediction result relative to the actual value reached
up to 43%, indicating that the selected features are not optimal for estimating the emission
rates of a given object.

A prerequisite for the methods described in this section to work is to acquire accurately
the value of minimum detectable concentration, i.e., detection limit. However, this detection
limit of passive OGI varies in different reports. Blinke et al. [155] reported that an ordinary
IR camera could detect 1–10 g/h leakage of hydrocarbon, while Ravikumar et al. [156]
reported that at a distance of 3 m, the detection limit of methane was about 20 g/h with
a 90% confidence by using a FLIR GF-320 camera. Some other studies [157,158] reported
that if the natural gas emission rate is larger than 30 g/h (about 1.5 scfh [159]), the leakage can be detected. This puzzling situation regarding detection limits is due to the fact that they are affected by multiple factors, such as the ambiance, camera characteristics, and molecular spectral characteristics.

It is intuitive that environmental factors such as temperature contrast between the environment and plume [19,38], the imaging distance [156], and atmospheric dispersion of the emitted plume (e.g., wind speed) [160] have an impact on detection limit. The environmental effect can be regarded as uniformly affecting any species measurement. However, the impact of the camera and the particular emission spectrum on the detection limits of different molecules varies and can be optimized. The detection limit is affected by the type of the detector (thermal detector, photodetector), the dark noise in the pixels of the camera chip that essentially determines the dynamic range of the device, and the possible existence of Peltier cooling for dark noise suppression. Also, the material of the detector, such as HgCdTe or InSb, which determines the bandwidth of sensitivity of the camera, has a strong influence on minimum detection limits [4].

Another one is the effect of molecular spectral characteristics, as shown in Table 1, molecules, such as CO₂ and CH₄, have multiple and distinguishable characteristic bands of varying intensity due to their multiple rotational–vibrational states [161]. The selection of a band in measurement has a significant influence on the detection limit. For example, CH₄ has a strong signal in the middle wave infrared (MWIR) [4]; however, other hydrocarbons such as C₂H₆ also have a strong IR activity in this range, because this is due to the dynamics of the C-H bond. At the same time, CH₄ also has another characteristic band in long-wave infrared (LWIR) which is separated from other hydrocarbons, so it is easier to isolate a “clean” signal in this band, but the signal is weaker and of course easy to be disturbed by noise.

As a result of these considerations, the camera response and filter bandwidth have to be matched to the specific species that is to be measured [162]. Taking the work of Wu et al. [42] as an example, in order to quantify CO, they first selected the 4.6 μm band due to its strongest absorption signal; then, they selected a bandpass filter to isolate the interference of CO₂ and water; and, finally, a mid-infrared camera was selected according to the filtered band. As a consequence of various combinations, the detectable limit of different molecules varies, and of course, the same imaging device has different detection limits for different molecules. For example, as mentioned in [16], an IMSS spectral imager in a particular optical configuration had a detectable level of butane at 15 g/hr, but 36 g/hr for ethylene. The detection limit is, therefore, application-specific [23], which means that the emission factors that are used in the minimum-detectable concentration methods also need to be recalibrated for each particular application.

6. Summary, Future Prospects, and Conclusions

A summary of current technologies of passive infrared imaging of gases as well as of the methods for the extraction of quantitative information from such data is given schematically in Figure 16. This is divided into three parts, i.e., tasks, methods, and requirements. Tasks include three core quantification tasks, namely, column density, concentration, and emission rate. The methods part includes the techniques used to realize the tasks by IOGI, while the requirements are the backbone technologies/hardware/principles supporting the development and application of these methods.

According to the number of times a particular requirement relates to a method, the rank of the importance of requirements is as follows, with the arbitrary assumption that each link appearing in the Figure is equally important: IR camera (6 + 2 links in the figure), prior knowledge about the particular application (4), application of artificial intelligence (4), conventional image processing (3), IR spectral imager (2), spectroscopy analysis (2), and classical LDAR methods (2).

Such an analysis would indicate the capabilities of the IR camera as the most important tool for extracting data that can then be processed in order to extract quantitative
information. In regard to the links highlighting this importance, we should probably also add the ones relating to spectral imagers equipped with filters or dispersing elements.

Figure 16. The relation map of passive OGI quantification methods.

The “second” most important aspects in this ranking are prior knowledge and artificial intelligence, notably of equal significance. By prior knowledge, we refer to the temperature of the emission cloud and the background, exhaust geometry, wind speed, etc. The extensive requirement of prior knowledge indicates that high-quality IR camera hardware is not sufficient for high-fidelity quantitative information, which, in turn, constitutes the main difficulty of realizing emission quantification via passive IOGI.

Artificial intelligence in IOGI can help alleviate the complexity of extracting quantitative data from IOGI. For instance, it can solve the problem of extraction of quantitative information, e.g., through machine-learning-based methods, when used in the augmentation methods of column density and concentration quantification. The significance of AI exceeds that of conventional image processing technologies and offers potent alternatives, such as AI-driven optical flow and 3D reconstruction. If one considers the links between AI and image processing, there is a total of seven links, which shows that algorithms are the vital component needed to realize quantitative IOGI.

IR spectral imager and spectroscopy analysis technology pair hardware and software and, although they do not have a large number of links, we should keep in mind that without them, augmentation methods cannot work. With this analysis, we do not mean to underestimate the importance of the classical LDAR and spectroscopy methods that are currently the gold standard in industrial practice, but rather to point to the potential of combining IOGI with high-quality IR imaging and AI-driven modeling and analysis. In
fact, in order to further promote the development of IOGI, it is necessary to learn from the mature and diverse results of LDAR.

Figure 16 appears to imply a hierarchical relationship between the three main tasks that IOGI can deliver in that the solution of column density supports the solution of concentration and then the solutions for concentration/column density support the solution for emission rate estimation. However, this hierarchical relationship is not needed since machine learning and minimum-detectable-concentration-based methods can be applied independently of this hierarchy.

Much as the extraction of quantitative information from IOGI is a formidable task, there are three specific aspects that can make the case for increased relevance in the near-term future of engineering practice:

1. The advent of carbonless fuels and fuels of reduced carbon trace. It is serendipitous but also a matter of fact that such fuels (H₂, NH₃, light alcohols, and oxymethylethers), which are increasingly considered indispensable parts of the energy portfolio during the rapidly emerging energy transition, will generate flames and plumes that will have substantially less complicated spectroscopy, mainly due to the complete lack or the substantial decrease of soot and carbon oxides. The related fields of infrared emission can be very reasonably expected to be much simpler in terms of quantitative interpretation.

2. The rapid progress in the field of AI algorithms, some of which can be transferred, as we showed above from other applications to IOGI. It is true that this explosive AI progress is on occasion non-uniform. For example, the machine learning algorithms applied in augmentation methods are still under development, while the machine learning algorithms used in 3D reconstruction or optical flow are very mature. It is our expectation that AI algorithms will be able to break through the constraints of their application specificity and heavy reliance on data. Transfer learning and meta-learning offer substantial hope on this front.

3. The emergence of the “physics-guided” and “physics-discovered” AI [163–166]. Powerful algorithms directly transferred from the field of computer science can be strengthened substantially if the physics underlying the acquisition of IOGI data is coupled with the mathematics of machine learning. Currently, algorithms rely to a substantial extent on prior knowledge of several features of the solution, shown in this review. These requirements could possibly be alleviated if, e.g., the mathematics of the reactive flow could be embedded into the models or learnt from the data. Meanwhile, it is exciting that machine learning models can provide physical insights, which will, in turn, stimulate further developments in theory and the systems themselves, e.g., through the application of symbolic regression [167,168]. Further improvements in terms of IR-imaging hardware, such as sensitivity in the near-IR and availability of relatively cheap band-pass filters, will also provide a boost to IOGI technologies.

The ultimate target is to develop technologies and systems of the accuracy, robustness, convenience, and affordable cost that will make them appropriate for engineering practice in a manner that will expand the capabilities of the currently used gas-analysis and spectroscopic techniques. This will require interdisciplinary research that will combine the necessary advances in terms of both reactive fluid mechanics and data science, extend 2D IR images to 3D spectral ones, coupled with spectroscopy, and utilize global structure information via AI, in order to acquire column density, concentration, and emission rate information from IR images.


26. Teledyne FLIR. Remote Tank Level Monitoring and Gas Detection with AI|Optical Gas Imaging|FLIR. Available online: https://www.youtube.com/watch?v=oNDg-cNTgMU (accessed on 15 February 2022).


88. Gupta, D.; Barlow, B.; Hugenholz, C.; Funk, W.; Ravikumar, A. A Field Trial of New Methane Detection Technologies: Results from the Alberta Methane Field Challenge. ESNT—Eng. 2020. in prep. [CrossRef]


100. Rossi, L.; Molinier, T.; Akhloufi, M.; Tison, Y. On the use of stereovision to develop a novel instrumentation system to extract geometric fire fronts characteristics. Fire Saf. J. 2011, 46, 9–20. [CrossRef]


110. Tancin, R.J.; Spearrin, R.M.; Goldenstein, C.S. 2D mid-infrared laser-absorption imaging for tomographic reconstruction of temperature and carbon monoxide in laminar flames. *Opt. Express* 2019, 27, 14184. [CrossRef] 
118. Huang, J.; Liu, H.; Cai, W. Online in situ prediction of 3-D flame evolution from its history 2-D projections via deep learning. *J. Fluid Mech.* 2019, 875, R2. [CrossRef] 
129. Golgicyaz, S.; Talu, M.F.; Onat, C. Artificial neural network regression model to predict flue gas temperature and emissions with the spectral norm of flame image. *Fuel* 2019, 255, 115827. [CrossRef] 


