Cathodoluminescence nano-characterization of semiconductors

Paul R. Edwards and Robert W. Martin
Department of Physics, SUPA, University of Strathclyde, 107 Rottenrow, Glasgow G4 0NG, UK

Email: paul.edwards@strath.ac.uk, r.w.martin@strath.ac.uk

Abstract. We give an overview of the use of cathodoluminescence (CL) in the scanning electron microscope (SEM) for the nano-scale characterization of semiconducting materials and devices. We discuss the technical aspects of the measurement, such as factors limiting the spatial resolution and design considerations for efficient collection optics. The advantages of more recent developments in the technique are outlined, including the use of the hyperspectral imaging mode and the combination of CL and other SEM-based measurements. We illustrate these points with examples from our own experience of designing and constructing CL systems and applying the technique to the characterization of III-nitride materials and nanostructures.

PACS: 78.60.Hk, 68.37.Hk

1. Introduction

Cathodoluminescence (CL) is the phenomenon of light emission from a material under excitation by an energetic electron beam. The use of CL as an image-forming mode in scanning electron microscopy (SEM) can be traced back to work carried out on the first, pre-commercial, SEMs [1]. While early applications were mainly restricted to the study of cathode ray tube phosphors and geological samples, the technique began to be used for the analysis of epitaxial semiconducting materials during the 1960s [2]. The development of the technique as a characterization tool was driven by a number of differences between CL and the analogous techniques of photo- and electroluminescence (PL, EL). The principal advantage of CL lies in the fact that the spatial resolution is determined by the distribution of excess carriers in the material, and is not therefore limited by diffraction in the collection and excitation optics inherent in all far-field techniques. This has allowed the technique to be pushed beyond the micron scale, and this paper will discuss its application to the nanometre-scale characterization of semiconductor materials and devices.

The generation of electron-hole pairs in a semiconductor under irradiation by energetic electrons occurs through the mechanism of impact ionization. Since the electron beam energy $E_{\text{beam}}$ used in the SEM will be typically between 3 and 4 orders of magnitude higher than the bandgap energy $E_g$ of the semiconductor, this will result in a carrier generation rate far in excess of the incident electron flux. Also, while not allowing selective excitation of transitions below a given energy, as in PL excitation spectroscopy, this feature does allow all transitions to be excited even for the widest bandgap materials such as diamond and AlN.

One further dissimilarity between CL, PL and EL is the potential for very different carrier injection levels. The small volume in which carriers are generated in CL can result in the carrier density exceeding the equilibrium levels in the semiconductor even at relatively low
beam currents. This provides a useful tool in CL characterization, with the current-dependence of the emission spectrum providing information on effects such as band-filling, for example. However, it also presents a possible pitfall when the effect is not taken into consideration, and care is not taken to maintain low injection conditions.

The spectral, spatial, temporal, angular and polarization-dependent properties of cathodoluminescence are all routinely resolved, with the results being interpretable along broadly similar lines to those from other luminescence techniques. Emission lines associated with bound excitonic states, donor-acceptor pair bands and defect-related features typically dominate the spectra, with variations due to changes in (amongst other things) alloying, strain, doping and carrier concentration all providing insights into the material properties. CL images show the spatial distribution of these phenomena, while also allowing the distribution of defects to be visualized due to their higher rates of non-radiative recombination, and allowing micro- and nano-scale optoelectronic devices to be directly imaged.

2. Technical aspects
2.1. Spatial resolution
The maximum spatial resolution achievable with CL is of particular interest, as it is this quantity which gives CL its prime advantage over alternative luminescence techniques. Factors potentially influencing the resolution include: the width of electron beam when it hits the sample surface; the volume of sample within which electron hole pairs are generated; and the subsequent movement of these charge carriers prior to their radiative recombination.

2.1.1 Beam size. While the minimum spot size achievable in modern field emission gun SEMs (FEGSEMs) now approaches the 1 nm level, this value will vary significantly with accelerating voltage, beam current and working distance. The spot will be largest for low voltages, for high beam currents and long working distances, which are unfortunately the conditions most desirable to achieve a high resolution, high intensity CL signal while leaving sufficient clearance for a light collector. The difficulty of achieving a small spot at low voltage can now be helped by the use of aberration correction in the objective lens, which has now been implemented in an SEM [3]. However, for machines with uncorrected lenses and thermionic tungsten filament sources, the spot size may still be the limiting factor in pushing to resolutions below a few 100s of nanometres.

2.1.2 Interaction volume. Generally more significant than the spot size is the lateral spread of the beam during its interaction with the sample bulk. Such interactions are commonly modelled using Monte Carlo electron trajectory simulations (e.g. [4-5]). Figure 1 shows the results of simulating the trajectories of 200 electrons impinging on a 1 nm spot on the surface of a gallium nitride (GaN) sample, repeated for accelerating voltages of 1, 3 and 10 kV. This shows the strongly non-linear dependence of the interaction volume on the incident electron energy for a thick sample. It also illustrates the inherent link between the lateral resolution and probe depth, with these two quantities being approximately equal. This observation is important when pushing the technique towards the nanoscale: it is clearly not possible to achieve nm-scale resolution without limiting the probed volume to the near-surface region, which in turn will generally not be characteristic of the bulk material. Although this may be disadvantageous in some cases, it can also have advantages when deliberately probing near-surface features such as quantum dots and wells [6], or mapping monolayer crystal growth islands (in which the surface sensitivity allows essentially atomic scale depth resolution [7]). If cathodoluminescence is instead carried out on a thin electron-transparent foil in a scanning transmission electron microscope (STEM) geometry [8] (in either an SEM or TEM instrument), then beam spreading within the sample becomes negligible for high accelerating voltages, and this mechanism no longer presents a practical limit to the spatial resolution achievable.
2.1.3 Carrier diffusion. Since the CL signal results from the recombination of charge carriers, it is clear that the resolution of the technique will not depend directly on the spatial distribution of their generation, but on their steady state distribution after carrier transport has been taken into consideration. Within high-purity semiconductors, it might be thought that the relatively long diffusion lengths compared with the dimensions of the carrier generation function would result in diffusion becoming the dominant factor in limiting the resolution. However, experimentally-confirmed calculations by Donolato [9] suggest that diffusion does not alter the carrier distribution under electron beam excitation sufficiently to significantly affect the image resolution. By numerically simulating the effect of drift from a generation volume approximated by a uniform sphere, it was found that the resolution of the electron beam-induced current (EBIC) technique in imaging localized defects was dependent primarily on the carrier generation profile, even when assuming an infinite diffusion length, L. This is due to the carrier density decreasing in the 3-dimensional case as \((1/r)\exp(-r/L)\) with distance \(r\) from a generation source, rather than as the more familiar 1-dimensional form \(\exp(-r/L)\). This conclusion can be applied equally to CL, which is similar to EBIC in depending on the excess carrier concentration profile for its spatial resolution. While diffusion lengths have been inferred from CL intensity profiles through extended defects [10], from voltage-dependent CL in layered structures [11], and by varying the distance between carrier generation and light collection [12], the effect on image resolution is small when compared with the generation volume.

Another consideration for the role of diffusion in CL resolution is that the diffusion length is a property of the bulk material. Most contrast features in CL, however, must by their nature result in a locally reduced diffusion length: for example, if carriers are more likely to recombine radiatively at a given point then their ability to diffuse away from this point will diminish. In this way, what is being measured with CL is at least partly the variation in carrier diffusion itself. These variations may be due to increased rates of non-radiative recombination (such as at dislocations or other defects), or potential barriers (such as at hetero-interfaces and QWs). The advantage of this for nano-CL is that features several orders of magnitude smaller than the diffusion length have been observed, e.g. resolving 20-nm-spaced QWs in GaAs with a bulk diffusion length of 10’s of µm [13], and measuring peak shifts in InGaN/GaN QWs over a ~10 nm distance [14].

2.2 Types of CL measurement
The simplest form of CL imaging is the panchromatic mode, in which no spectral selectivity is applied. As the electron beam scans across the sample, all the collected light is directed to a single detector, producing a greyscale image whose intensity is the product of the emission
Cathodoluminescence nano-characterization of semiconductors

spectrum and the spectral response of the system. The monochromatic mode extends this by imaging only a single band of wavelengths, with the spectral selectivity being provided by a band-pass filter or spectrometer. In the latter case, the same CL system can be used to carry out spot-mode spectroscopy, with the spectrometer stepping through a range of wavelengths while the electron beam remains fixed. More recently, parallel detection using an array sensor has become a more common method of spectroscopy, although this can impose tighter constraints on the optical design, as will be discussed later.

Christen et al. [7] took advantage of improvements in multichannel detector design and computer processing power to develop a CL system capable of recording a full luminescence spectrum at each point during a beam scan. The technique is now widely used, and is variously referred to as CL spectrum-imaging, CL wavelength imaging (CLWI) or CL hyperspectral imaging (a term borrowed from the field of remote sensing, in which it is defined as an image resolved into 20 or more contiguous wavelength channels). The resultant multidimensional dataset must be further processed to produce spectra, linescans or 2-D images representing subsets of the data. For example, intensity images can be calculated over given wavelength ranges, giving similar results to monochromatic imaging. The most powerful aspect of the technique, however, is its use in mapping other characteristics of the luminescence which cannot be probed using the monochromatic approach, such as peak energies and widths. This can be achieved using peak fitting, in which a non-linear least-squares fit is made to each spectrum in the dataset, and the resultant fit parameters then used to build up a 2-D image [6].

A further advantage of the hyperspectral imaging mode lies in the possibility of applying multivariate statistical analysis techniques to the multidimensional datasets. Techniques such as principal component analysis and factor analysis have been applied to CL data from geological samples [15] and plasmonic nanostructures [16], allowing the most significant spectral components to be automatically deconvolved and mapped.

2.3. Optical design considerations

The idea that it is always preferable to collect CL over a maximum possible solid angle is misconceived, and there are a number of reasons why it may be neither possible nor preferable to attempt to intercept rays over the full $2\pi$ steradians above the sample surface. We will consider the basic challenge – to couple as much light as possible from the beam/sample interaction volume to the detector – together with the specific requirements of CL measurements and some basic optical design principles.

The optical design requirements for panchromatic imaging are modest, with a photomultiplier tube or photodiode offering a large detector area. Early CL optics designed for use in this mode utilized a simple light pipe arrangement [17], and a greater signal could be obtained by collecting over a larger solid angle, for example using a parabolic mirror [18] to reflect the collected light to the large detector. Moving to spectroscopy or monochromatic imaging placed slightly stricter requirements on the optical design, with the collected luminescence now needing to be brought to approximate focus at the entrance slit of a spectrometer whilst matching the $f$-number of the entrance aperture. However, wide slits could be used, while still maintaining high spectral resolution by compensating with a higher density ruled grating. In hyperspectral imaging mode, this wide slit / fine grating combination comes at a cost to the observable spectral range and more careful optical design now becomes essential.

Efficient coupling of the CL into a spectrometer represents a challenge similar to that faced in all luminescence measurements. However, use of a scanning luminescence spot gives the additional problem of collecting light from an extended source into a small detector slit. The problem faced can be described using the concept of étendue, a quantity which is invariant through an optical system and which can be approximated as the product of the source area and the solid angle of the entrance pupil.
Figure 2 shows a simple imaging system which illustrates the effect of étendue in limiting the field of view (FOV) of CL collection optics. The optics collect luminescence from the sample with a half-angle $u$, fixed by the numerical aperture of the objective (or collection solid angle). This light is refocused to form an image at the spectrometer slits with a half-angle $u'$, which must be matched to the $f$/number of the instrument used. These two angles thereby fix the magnification of the collection system, irrespective of the complexity of any intermediate optics. The scanned region of size $\eta$ is therefore imaged to the slits with an image size given by:

$$\eta' = \eta \frac{\sin u}{\sin u'}$$  \hspace{1cm} (1)

![Diagram of optical system with labels](image)

**Figure 2: Simplified optical design schematic for luminescence collection.**

If collection optics with greater solid angle are used, the magnification of the system increases, and the FOV (limited by the spectrometer slit in the image plane) will reduce further. This can be calculated from the simple geometric expression:

$$\text{FOV} = \frac{d}{2 \cdot (f/\#) \cdot \tan u}$$  \hspace{1cm} (2)

where $d$ is the spectrograph slit width. If we assume the angular-dependence of the emission to follow a Lambertian profile (i.e. $\propto \cos \theta$), and the collection axis to be normal to the sample surface, the fraction of emitted light which is collected is found by integrating $\cos \theta$ over the collection cone:

$$\text{collection} = \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} \cos \theta \sin \theta \, d\theta d\phi = \sin^2 u$$  \hspace{1cm} (3)

Using these expressions, the calculated curve in Figure 3 shows the FOV for different collection solid angles in a typical CCD spectrograph CL system. This assumes a fast spectrograph ($f/4$), and a slit width $d$ optimized for a typical modern sensor (25 $\mu$m, corresponding to the pixel pitch of a 1000-channel 1-inch CCD chip).
Cathodoluminescence nano-characterization of semiconductors

Figure 3: Maximum CL collection efficiency versus field of view, calculated for an f/4 spectrometer with 25 µm slits.

This plot shows for example that if a 10 µm field of view is required, no more than 10% of the emitted light could be coupled into the spectrograph without overfilling the grating. This calculation represents an upper limit to the maximum collection efficiency possible with this spectrograph/slit combination; the substitution of higher f-number spectographs, off-axis collection (as is the case in many CL systems) or a hole for the electron beam will all further decrease the maximum possible collection for a given field of view. The same argument also holds for other collection geometries, such as the commonly-used conic (parabolic [18] or elliptical [19]) collecting mirrors and fibre-coupled systems.

This problem could be addressed by the use of “de-scanning” mirrors in the optical path, to produce an in situ scanning optical microscope with a focus that follows the scanning electron beam. A simpler solution is the use of a fixed beam and scanning stage, especially where such a facility is already built into the instrument, as is the case in electron probe micro-analyzers (EPMAs) [20].

An additional consideration is that the versatility of scanning electron microscopy is due in large part to its multimode functionality, with different beam-induced signals being measured simultaneously during a scan. This allows different properties of a material to be probed on a sub-µm scale without having to relocate and align the same sample area. There are many cases in semiconductor characterization where it is desirable to measure cathodoluminescence together with another SEM signal. Examples include: CL and basic secondary and backscattered electron signals for topographic imaging; CL and energy/wavelength dispersive X-ray (E/WDX) spectroscopy [20-21] for probing the relationship between composition and emission; CL and EBIC [22] for the comparison of radiative and non-radiative recombination characteristics; and CL with electron backscattered diffraction (EBSD) to investigate the effect of strain on luminescence [23]. In the majority of cases, collection optics which attempt to cover most of the available 2π steradians will be incompatible with these other measurements.

For all these reasons, the CL systems designed and built by the authors have avoided the use of large solid angle conic mirrors in favour of lower-N.A. optics with high efficiency, f-matched, direct coupling to the spectrograph.
3. Examples

3.1. Combined CL and X-ray fluorescence

The luminescence properties of optoelectronic materials are strongly dependent on their elemental composition, whether this be the precise alloy within a multinary system, the degree of stoichiometry within a compound semiconductor, or the levels of gross contaminants such as oxygen. Even within unpatterned bulk materials, these properties can all vary on a microscopic level, and will accordingly influence the emission properties of the material on a similar length scale. Exploiting the multimode nature of SEM, it is often useful to carry out cathodoluminescence together with composition measurements based on X-ray fluorescence generation, such as EDX or WDX spectroscopy. These techniques are well matched to CL, with approximately similar volumes of material being probed in each case, and simultaneous measurement allows nano-scale variations in each to be mapped without the alignment issues associated with data acquired ex-situ.

Figure 4 shows an example of measuring WDX and hyperspectral CL images simultaneously from a large sample of the quaternary alloy Al\textsubscript{x}In\textsubscript{y}Ga\textsubscript{1-x-y}N [21]. The instrument used was a modified Cameca SX100 electron probe microanalyzer equipped with multiple WDX detectors, in which the CL is collected using an optical microscope coaxial and confocal with the electron beam. The signals were detected with a stationary electron beam and scanning sample stage. In this small subset of the total data acquired, a clear correlation is seen between the maps of the emission energy and the fraction of InN in the alloy. This connection can be seen graphically in the correlation plot of Fig. 4(d), a 2-dimensional histogram in which values from corresponding pixels in the two images are plotted against each other. Such analysis is equally applicable to higher resolution data, such as in the study of sub-µm variations of composition and luminescence in In\textsubscript{x}Ga\textsubscript{1-x}N epilayers [24].

![Figure 4: Data from an AlInGaN sample, showing maps of (a) InN fraction and (b) peak CL energy, (c) example spectra from the CL dataset and (d) a correlation plot of CL energy vs. InN% generated from the two maps.](image)

3.2. CL spectroscopy at short wavelengths

Figure 5 shows room temperature CL spectra from a series of Al\textsubscript{x}In\textsubscript{y}N, epilayers grown with different composition values \(x\) on GaN buffer layers [25]. The spectra show composition-dependent AlInN near-band-edge luminescence peaks in the 300-350 nm range, and a second peak at 363 nm emitted from the underlying GaN. This data was acquired from a FEGSEM-based CL system which utilizes a Schwarzschild-type objective to collect the emission from a sample inclined at 45°, and demonstrates an advantage of using an all-reflecting optical design. In the near-UV (~400-300 nm) and middle-UV (~300-200 nm) spectral ranges, refracting optics tend to suffer from high absorption and also excessive chromatic aberration due to the strong dispersion of optical materials at these wavelengths. While expensive silica/CaF\textsubscript{2} UV achromats could be used to cover this region, the optics would then not be...
Cathodoluminescence nano-characterization of semiconductors

corrected for use in the visible region. This system has also been used to examine even wider bandgap materials including Al\textsubscript{x}Ga\textsubscript{1-x}N [26] and diamond [27].

![Normalized CL spectra for different Al\textsubscript{x}In\textsubscript{1-x}N alloy compositions](image)

**Figure 5:** Normalized CL spectra for different Al\textsubscript{x}In\textsubscript{1-x}N alloy compositions, demonstrating the short-wavelength capability of an all-reflecting optical design.

### 3.3. CL of nm-scale features

The same CL system was used to acquire the results shown in Figure 6, from CL hyperspectral imaging of an array of nano-scale GaN pyramid structures containing an InGaN/GaN QW [28]. Peak fitting to the QW exciton emission has been carried out to generate the image in Fig. 6(b), and Fig. 6(c) shows a linescan through this image. This data shows shifts in the emission peak energy taking place over length scales of only a few 10s of nm, which in this case is interpreted as an effect of varying InN incorporation in the well.
Cathodoluminescence nano-characterization of semiconductors

![Figure 6](image)

Figure 6: (a) SEM image of GaN nano-pyramids containing a single InGaN QW, (b) map of fitted CL peak energy, and (c) linescan through this image.

3.4. CL collected using different geometries
The angular emission profile of the luminescence from a planar surface is often assumed to be Lambertian, with intensity dropping off as the cosine of the angle from the surface normal. In some cases, however, the presence of surface morphology or sub-surface structure may result in a more complex directionality to the emission. In addition, it can also be the case that not only the intensity but also the spectrum of the emission varies as a function of angle. Such structures include, for example, photonic crystals and vertical cavity devices, and this effect must be considered when measuring CL from these samples. As an example, Figure 7 shows a schematic of a laser microcavity device containing In$_{x}$Ga$_{1-x}$N/GaN quantum wells, mounted for cross-sectional CL analysis and showing the collection cone corresponding to the N.A. 0.28 objective used.
Rotating the sample eucentrically about the beam axis changes the path of the collected light, from direct emission to luminescence filtered through one of the cavity mirrors. The effect on the recorded spectrum can be seen in Figure 8, in which the full-width at half maximum of the cavity-filtered emission is reduced compared with the directly-measured peak, from 120 nm to 50 nm. Of course, this very marked change represents an extreme case; nonetheless, while the effect will not be as pronounced in less deliberately directional structures, it should still be considered. A more subtle example of the influence of sample geometry, for example, would be the effect of self-absorption of near-band-edge luminescence on a sample with strong surface morphology. In this case, luminescence rays which pass through a greater thickness of material before collection will produce peaks slightly narrowed and red-shifted, which can lead to misinterpretation of a hyperspectral image.

Figure 8: CL spectra from an InGaN/GaN quantum well microcavity, showing the effect on the observed spectrum of varying the collection geometry.

4. Conclusions
While cathodoluminescence is a well-established method in the field of semiconductor microcharacterization, ongoing improvements in microscopes, detectors and data analysis methods have ensured that the technique has continued to develop. With well-designed optics and careful consideration of the beam-sample interaction, it is a technique which can be extended well into the nano-scale regime.
Cathodoluminescence nano-characterization of semiconductors

Acknowledgments
We would like to thank Nils Brauer and Nicolas Grandjean of the École polytechnique fédérale de Lausanne for supplying the InGaN/GaN microcavity and for the inclusion of unpublished spectra from this sample. We are also grateful to our collaborators who provided samples for the other results reproduced in this work: I. M. Watson (University of Strathclyde); C. Liu, P. Shields, D. Allsopp and W. Wang (University of Bath); and S. Fernández-Garrido and E. Calleja (Universidad Politécnica de Madrid). This work was supported through a UK EPSRC platform grant.

References