REVIEWS OF TOPICAL PROBLEMS

PACS numbers: 07.79.-v, 61.16.Ch

Cathodoluminescence microscopy

V I Petrov

Contents

1.	Introduction	807
2.	Formation of a cathodoluminescence radiation	808
	2.1 Generation of nonequilibrium charge carriers; 2.2 Nonequilibrium carrier motion and recombination	
3.	Spatial resolution of a cathodoluminescence scanning microscopy	810
4.	Instrumentation	812
5.	Information obtained from a cathodoluminescence signal	813
	5.1 Integral cathodoluminescence intensity; 5.2 Cathodoluminescence spectra; 5.3 Polarization	
	of cathodoluminescence radiation	
6.	Perspectives of the technique development and conclusions	817
	References	817

Abstract. Being the source of light quanta in optoelectronic devices like semiconductor lasers and light-emitting diodes, radiation recombination processes are of great importance in the constituent semiconducting structures. The latter consist of thin epitaxial layers, and their radiative recombination efficiency depends on structural imperfections such as dislocations, usually acting as nonradiative recombination centres, and point defects. Since these imperfections usually have a nonuniform layer distribution, radiation characteristics are distributed inhomogeneously over the layers and small-scale structural elements. For such structures, therefore, cathodoluminescence scanning microscopy, controlling the microlevel distribution of radiation characteristics, is an irreplaceable research tool. The effective use of this technique requires a correct understanding of the radiation-producing processes occurring during the electron probe-semiconductor interaction. A knowledge of the potential and the information capacity of the method is essential.

1. Introduction

A cathodoluminescence (CL) phenomenon, i.e. emission of light under electron bombardment, is known for a long time and now is widely used in nearly all black-and-white, and colour cathode-ray tubes. Moreover by means of a cathodoluminescence the luminescent properties of different materials are studied as well. Extensive references on CL investigations can be found in monographs [1, 2], and more later works up to 1988 were partially summarized in Ref. [3].

V I Petrov M V Lomonosov Moscow State University Vorob'evy Gory, 119889 Moscow, Russia Tel. (7-095) 939-14 48 E-mail: vtim@ofme.phys.msu.su

Received 20 December 1995 Uspekhi Fizicheskikh Nauk **166** (8) 859–871 (1996) Translated by V I Petrov

When studying different objects the CL technique comes out in the same row together with other spectroscopic methods such as a photoluminescence and so on. So far as mechanisms leading to the emission of light in a semiconductor and connected with nonequilibrium charge carriers recombination are analogous for different forms of excitation, the results obtained will be similar, however, some differences may exist. The electron beam excitation, for example, leads to the light emission by all mechanisms of radiative recombination which are present in a semiconductor. By varying the electron-probe energy one can obtain depth-resolved information about the object due to a changing of electron penetration depth into the object. There is also a more essential difference between light and electron excitations as it has been shown in Ref. [4]. In the latter case, as far as electrons possess also a momentum in addition to their energy, the nonequilibrium carriers appear in indirect bands as well, and supplementary peaks in CL-spectra (e.g., peak S in Fig. 1) arise.

The development of optoelectronics stimulated the elaboration of methods which enable us to control local parameters of micron and submicron layers in semiconductor epitaxial structures with high spatial resolution. One of such methods is a local cathodoluminescence or microcathodoluminescence — the cathodoluminescence mode of a scanning electron microscope (SEM), which permits a control of the luminescent parameters distribution over the object surface to be established. Here the CL radiation is excited by a thin electron probe focused at a given point of the object. Besides the high spatial resolution obtained in this case, the microscope provides an additional information about the object, such as a surface topography, chemical composition, and presence of electrical barriers, if secondary electron mode, Xray microanalysis mode and electron-beam-induced current (EBIC) mode are also used.

It should be noted that at present nearly all the SEM produced in Japan, USA, England, France and some other countries, are supplied with standard CL attachments and liquid-helium cooling stages, mostly manufactured by Oxford Instruments, as well as with an attachment for EBIC



Figure 1. The photoluminescence (above) and cathodoluminescence (below) spectra from the $GaAs/Al_xGa_{1-x}As$ sample with multiple quantum wells at 10 K [4b].

measurements. The latter is very important for studying the semiconductors so far as the presence of EBIC-signal indicates the electrical activity of the object area under investigation (an inhomogeneity).

2. Formation of a cathodoluminescence radiation

The three fundamental processes resulted in a CL emission are the generation, motion, and recombination of excess electrons and holes. Let us consider these processes individually.

2.1 Generation of nonequilibrium charge carriers

Primary (i.e. incident or probe) fast electrons striking the object undergo the energy losses due to interaction with solid, and then stop. In order to evaluate these energy losses the model of a continuous energy transfer is usually used for simplicity. The mean rate of electron energy loss per unit length of its path s in the solid (all dissipation mechanisms being included) can be described by the Bethe expression [5] as

$$\frac{\mathrm{d}E}{\mathrm{d}s} = -2\pi e^4 N_\mathrm{A} \, \frac{\rho Z}{EA} \ln \frac{1.166E}{J} \,, \tag{1}$$

where *e* is the electronic charge, N_A is Avogadro's number, ρ is the density of the object material, *A* is the atomic weight, *Z* is the atomic number, *E* is the mean electron energy, and *J* is the mean ionization potential.

The different secondary processes caused by electron bombardment proceed in the energy dissipation area, socalled excitation volume, because of electronic transitions in the material. The dimensions of this excitation volume are close to the value of penetration depth or range of primary electrons R_e which, according to Ref. [6], is

$$R_{\rm e} = \frac{0.0276A}{\rho Z^{0.889}} E_0^{1.67} \, \left[\mu {\rm m}\right],\tag{2}$$

where A is in g mol⁻¹, ρ is in g cm⁻³, and E_0 is the electron probe energy in keV. Figure 2 shows the electron penetration depth calculated, according to the Kanaya-Okayama model [6], for several materials.



Figure 2. The depth R_e of primary electron penetration into the object as a function of electron energy E_0 calculated according to formula (2).

The shape of the excitation volume depends on the atomic number: it varies from a pear shape for a low-atomic-number material, to a near spherical shape for 15 < Z < 40, and to hemispherical one for larger-atomic-number materials. It is worth marking the difference in the excitation volume shapes in the thin film and bulk objects (see Fig. 3).



Figure 3. Excitation volumes in the thin and bulk objects.

The dimensions of the area where one of the secondary processes takes place, otherwise the generation area for the accompanying signal, are conditioned by the activating energy of this process. In semiconductors, the energy of incident electrons dissipated in the object may be partially spent on the creation of nonequilibrium electron-hole pairs and this causes an essential increase in the number of the mobile charge carriers. The valence electrons can move from different levels of valence band to these of the conduction band, but laying higher than the maximum level of electrons created by thermal generation. That is why the average energy required for the generation of one electron-hole pair E_i is somewhat larger than the band-gap energy E_g , and equals to

$$E_{\rm i} = 2.8E_{\rm g} + M\,,\tag{3}$$

with 0 < M < 1 eV depending on the material and being independent of the electron probe energy. The values of E_i for several materials of interest are given beneath (those of E_g at 300 K are given in parentheses): InSb — 0.42 eV (0.16 eV), PbS — 2.05 eV (0.41 eV), GaAs — 4.6 eV (1.43 eV), CdTe — 4.65 eV (1.5 eV), GaP — 7.8 eV (2.24 eV).

In the SEM CL mode the energy of an incident electron is essentially larger than the average energy of an electron-hole pair production, and one 10-20 keV electron can generate several thousands of electron-hole pairs along its trajectory in the excitation volume. Their number is a generation factor given by

$$G = \frac{E_0(1-\gamma)}{E_{\rm i}} , \qquad (4)$$

where γE_0 represents the beam energy loss due to an electron backscattering from the object.

Besides the electron-hole pair production, fast beam electrons can cause a local heating and changing of structure and properties in the irradiated area, especially in the focused mode of the SEM. It may lead, for example, to the change in the local luminescent properties (see Refs [7, 3]).

2.2 Nonequilibrium carrier motion and recombination

Nonequilibrium charge carriers generated in a semiconducting material acquire the directional motion by diffusion and drift processes, and recombine simultaneously because of the electron returning to the free states in valence band that results in free electrons and free holes disappearing. In thermodynamically equilibrium state the generation and recombination processes compensate each other completely. Depending on a mechanism three modes of recombination can be recognized — interband recombination, recombination via localized states in the gap and surface recombination. According to how the energy released in the course of recombination is spent on, the recombination may be radiative with photon emission as well as nonradiative with the energy transfer to the lattice (phonon emission), or to another electron which changes to a higher-energy state in the conduction band (impact recombination or Auger effect).

For a continuous irradiation and the electron beam at a fixed position on the object, the action of all these processes results in the formation of a stationary excess-minority carriers distribution which is governed by the three-dimensional differential equation of continuity. In the simplest case for electrons in a p-type semiconductor it can be written as

$$D\nabla^2 \Delta n(r) - \frac{\Delta n(r)}{\tau} + G(r) = 0$$
(5)

with a boundary condition at the object surface

$$D \left. \frac{\partial \Delta n(r)}{\partial z} \right|_{z=0} = v_{\rm s} \Delta n(r) \big|_{z=0} , \qquad (6)$$

where $\Delta n(r)$ is the excess-minority carrier density per unit volume, *D* is the diffusion coefficient, τ is the minority carrier

lifetime, v_s is the surface recombination velocity, G(r) is the generation function (source function) of electron-hole pairs defined by the electron beam parameters and the energy dissipation processes in the object or, in other words, the energy loss density distribution function in the object, and, finally, z is the coordinate from the surface into the object (see, for example, Ref. [8]). The equation (5) is valid under the following two conditions:

(i) The carrier lifetime τ is independent of Δn that is satisfied quite well as long as Δn is small compared to the majority carrier density, i.e. when $\Delta n \ll p_0$, where p_0 is the equilibrium hole concentration.

(ii) The motion of excess carriers is purely diffusive (internal electric fields are lacking in the object).

The intensity I_{CL} of the CL emission, which is usually accepted to be proportional to $\Delta n(r)$, can be written as

$$I_{\rm CL} = \int_{V} AB \eta(r) \,\Delta n(r) \,\mathrm{d}V_r \,, \tag{7}$$

where *A* and *B* are the correction factors for CL absorption losses inside the object, while the light is passing through the object material, and CL reflection losses at the object-vacuum boundary, respectively, and $\eta(r)$ is the internal quantum efficiency defined as the ratio of the radiative recombination rate to the total recombination rate, which is the sum of the nonradiative rates (from all sources) and the radiative rate. The internal quantum efficiency η is usually expressed in terms of carrier lifetimes – radiative τ_{rr} , and nonradiative τ_{nr} as

$$\eta = \frac{1}{\tau_{\rm rr}} \left(\frac{1}{\tau}\right)^{-1},$$

where

$$\frac{1}{\tau} = \frac{1}{\tau_{\rm rr}} + \frac{1}{\tau_{\rm nr}} ,$$

and thus

$$\eta = \left(1 + \frac{\tau_{\rm rr}}{\tau_{\rm nr}}\right)^{-1}.\tag{8}$$

To calculate I_{CL} from Eqn (7) we need to know the exact solution of Eqn (3). This is feasible only for homogeneous material, when the minority-carrier diffusion length $L = \sqrt{D\tau}$ is constant, and simple boundary conditions (4) valid for a point source or a spherical source with a constant pair density. Very often in calculations it is considered that $v_s = \infty$. These results can be regarded as true for arbitrary sources also, but only if distances between sources and boundaries are large as compared with L and generation volume size.

Besides the intensity, CL radiation is also characterized by its spectrum. The photon energies and, therefore, CL-radiation spectrum contain information about the characteristic energy levels present in the object. Radiative transitions can be divided into two classes: (i) intrinsic (fundamental), and (ii) extrinsic. Intrinsic transitions include interband transitions across the fundamental energy gap E_g with the emission of a photon of energy close to E_g , and a free exciton decay with the emission of a photon of energy $hv = E_g - E_x$, where E_x is the exciton ionization energy. In indirect-gap materials, such as, for example, GaP, the fundamental emission is very weak, since unlike direct-gap materials, such as GaAs, the recombination of electron-hole pairs must be accompanied by the simultaneous emission of a phonon, and the appearance of an extra particle decreases the probability of such a process significantly as compared with direct transitions (see Fig. 4). Therefore, the fundamental emission in indirect-gap semiconductors is relatively weak, especially when compared with that due to impurities and defects.



Figure 4. Direct (a) and indirect (b) interband transitions: E_1 and E_2 are the initial state and final state energies, E_{ph} is a photon energy.

The transitions between the levels inside the energy gap, associated with a presence of donors, acceptors, and native defects, are extrinsic in nature. Extrinsic levels may be classified as shallow, situated near the band edges, or deep, placed closer to the middle part of the gap. Transitions via shallow levels lead to recombination emission with photon energies close to E_g , similar transitions via deep levels can also lead to recombination emission but with photon energies well below E_g . Such extrinsic radiation can be made much more intense than intrinsic CL at ambient temperatures even in direct-gap materials. Using phosphor terminology, the luminescence emission in these cases is 'activated' by impurity atoms and other defects, so-called activators. In the presence of impurities, bound excitons may be formed, and after their recombination the photons with energies

$$hv = E_{\rm g} - E_{\rm x} - E_{\rm b} - mE_{\rm ph}$$

arise, where $E_{\rm b}$ is the binding energy of the exciton to the impurity atom, $E_{\rm ph}$ is the energy of the emitted phonons, and m is the number of the phonons emitted. In general, CL emission from real materials is dominated by extrinsic recombination mechanisms through the levels associated with impurity atoms and defects.

3. Spatial resolution of a cathodoluminescence scanning microscopy

Besides main spectroscopic features of the system intended to analyze the light radiation (such as a total spectral transmission characteristic of the system, a spectral resolution, and a detection sensitivity limit), the local CL technique is characterized by one more very important parameter, namely, a spatial resolution or a locality of the information collection, i.e. the minimum distance between two object details which we are possible to distinguish using the CL-signal. For bulk materials it is much worse than a spatial resolution of the SEM in the secondary electron mode, which usually characterizes a SEM. The reason is that the CL emission occurs from the volume much larger than one pertaining to secondary electron escape. It includes the immediate volume of pair generation under electron exposure and the space wherein excess-carriers diffuse after their creation.

Due to a smallness of the CL signal, especially for materials with a modest internal quantum efficiency, larger probe currents are required $(10^{-6} \text{ to } 10^{-8} \text{ A})$ for operation in the local CL mode. It results in a divergence of the probe electron beam at the object and makes an additional contribution to the spatial resolution degradation. Based on the early ideas, the authors of a review [8] suggested that the spatial resolution in a CL mode could be expressed as a geometrical sum of the beam diameter at the object, the lateral dimension of the energy dissipation volume (the excess-carrier generation volume), which was supposed to be equal to the electron penetration depth $R_{\rm e}$, and the minority charge-carrier diffusion length L. It should be noted that this standpoint was supported without any proof in more later works (see, for example, [9]). This estimation gave the spatial resolution value which was essentially larger than that realized in practice. For example, in [10] the resolution of 1.5 μ m has been reported for *n*-GaP with $L = 3.5 \mu$ m at 20 kV, instead of 6 µm as follows according to the calculation with the aid of the geometrical sum formula.

Later on, in Refs [11, 12] a formal theory of the spatial resolution of the local CL has been developed. This theory was founded on the definition of the spatial resolution in the CL mode relying in turn on the criterion of resolution of two points in the transmission electron microscopy (see [13] and [14]), where the dip between two maxima on the image of two points resolved was considered to reach 25%. In the local CL, the minimum distance between two thin luminescent vertical layers in a nonluminescent matrix was accepted as a spatial resolution limit when the dip between maxima in the CL signal (response), while the probe was scanned across these layers, ran to 25%.

It turned out that the spatial resolution defined in this way is very close to the half-width of the CL response to thin luminescent layer in the object, i.e. full width at half-height of the maximum of the CL signal when the beam is just on the luminescent layer. In theoretical consideration, contributions due to a scattering of primary electrons in an object (finite size of a pair generation volume) without an excess-carrier diffusion, and due to excess-carrier diffusion without scattering (neglecting a size of pair generation volume) have been allowed for separately. Influence of the finite size of the beam on the spatial resolution was neglected because of its small contribution.

The diffusion length L was assumed to be zero when only a finite size of a pair generation volume has been taken into account. The CL response, while scanning across a thin luminescent layer, has quite a narrow central part with a higher signal amplitude (Fig. 5), and its half-width determines the spatial resolution. In this case it makes a contribution to the spatial resolution due to a primary electron scattering in an object, and its value depends on the microscope accelerating voltage (Fig. 6). The shape of this dependence is unlike for the materials with various atomic numbers because of the difference in the scattering processes. A small-angle inelastic scattering dominates in light materials (Si), whereas a largeangle elastic scattering — in heavy materials (PbS, GaAs). The spatial resolution obtained in such calculations can be considered as a limiting spatial resolution, which is reached either on the materials with a small diffusion length or in a pulse mode of the SEM. In the latter case the registration of



Figure 5. Normalized local CL response signal $I_{CL}(x)$ while scanning across a thin luminescent layer in the GaAs matrix calculated by a Monte-Carlo method for different accelerating voltages: I = 10, 2 = 20, 3 = 30, 4 = 40, and 5 = 50 kV without an excess-carrier diffusion [11].



Figure 6. Full width at half-height δ of the CL response $I_{CL}(x)$ as a function of the accelerating voltage without an excess-carrier diffusion: I - Si, 2 - GaP, 3 - PbS and GaAs [11].

the CL signal is made only in the initial moment of the electron beam excitation as long as the essential diffusive divergence of nonequilibrium charge carriers would be occurred, as it was shown in [15]. The enhancement of the spatial resolution in this case is more noticeable, the larger the diffusion length (it has to be taken into account the necessary increase of signal recording time because of the essential decrease in its intensity). It should be also noted that a new experimental technique for evaluating the size of an electronhole generation volume has been proposed recently (see Ref. [16]). For this purpose an AlAs/GaAs structure with three quantum wells was used, and the CL signal was monitored while the probe was scanned across the layers. In such an object the charge carrier diffusion has been excluded, and the spatial resolution increased up to 50 nm. The CL responses, obtained by the authors, were similar to those in Fig. 5.

The calculation of analogous CL responses for a point pair generation source in the presence of carrier diffusion $(L \neq 0)$ showed a dependence of the distribution half-width on a surface recombination velocity, and this dependence became more pronounced when a point source approached the surface (Fig. 7). The presence of surface recombination leads to the carrier influx into the surface, where they nonradiatively recombine at a later time. At the expense of this effect the diffusive divergence of carriers as well as a CL response half-width decrease. In this case a total intensity of the CL signal decreases too. Such narrowing of the CL response curve due to the presence of surface recombination may amount to as low as one seventh for a bulk object and a point pair-generation source. Due to the presence of surface recombination the minority-carrier lifetime near the surface differs from that in the bulk, τ_{bulk} , and takes some effective value

$$\frac{1}{\tau_{\rm eff}} = \frac{1}{\tau_{\rm bulk}} + \frac{1}{\tau_{\rm s}} ,$$

where τ_s is the carrier lifetime due to recombination at the surface. The influence of surface recombination increases essentially when the objects are thin. Here the presence of



Figure 7. Normalized full width at half-height $\delta^* = \delta/L$ of the CL response $I_{\text{CL}}(x)$ as a function of the normalized surface recombination velocity $S = v_s \tau/L$ for different reduced point-source depth $Z_0 = z_0/L$: $I = Z_0 = 0, 2 = 0.05, 3 = 0.1, 4 = 0.2, 5 = 0.3, 6 = 0.5, 7 = 1.0$, and 8 = 2.0 [12].

already two surfaces bounding the object from above and below has to be taken into account. In this case

$$\tau_{\rm eff} = \frac{t^2}{\pi^2 D} \,,$$

where *t* is the object thickness. From above it is clear that even though no recombination occurs in the bulk, i.e. $\tau_{\text{bulk}} = \infty$, the lifetime will be only defined by the presence of surfaces, and the corresponding diffusion length will take the form

$$L_{\rm s}=\sqrt{D\tau_{\rm s}}=rac{t}{\pi}$$
 .

The effect of surface recombination was also recognized in [17] when the author worked with very low accelerating voltages (1 kV) in an effort to heighten this effect, and realized the record resolution of 60 nm in the CL mode. Another possible way to improve the resolution in a CL mode is the application of a magnetic field to the object which is directed along the electron beam. In this case the Lorentz force prevents excess carriers from being diffusively spreaded. For the first time it was suggested in [18], and much later in [19]. In scanning tunnelling microscope (STM), when using low bias voltages, it is possible to excite the luminescence in a small volume of the object as it has been done in [20] under spectroscopic investigation of the luminescence from a single quantum wire. However, here the mechanism of the luminescence is alternative: the electron-hole pair generation is lacking but holes are injected into the object, and then recombine with thermolized electrons. The luminescent spectra, obtained in the STM, are shown in Fig. 8.

With a decrease of the object thickness the size of the excess-carrier generation volume also reduces due to a smaller contribution of scattered primary electrons. The divergence b of the primary beam for a small object thickness t, small size



Figure 8. Luminescence spectra obtained in STM at two different tip positions over the GaAs/AlGaAs structure with a quantum wire array (with a period of 765 nm) at the surface: just over a quantum wire (I), and between two neighbouring wires (2). In curve (I) the peak is seen which corresponds to the luminescence from the quantum wire, in (2) — the luminescence from the quantum well layer at the object surface only [20].

of beam on the object (3 nm), and only elastic interaction of electrons with the object [13] is given by

$$b \approx \frac{Z}{E_0} \left(\frac{\rho}{A}\right)^{1/2} t^{3/2} \,. \tag{9}$$

The dependences of this quantity on the electron beam energy are shown in Fig. 9 for 1 and 2 μ m thick Si films. For 100 keV electrons, the beam divergence in the 1 μ m film is only 0.2 μ m, thus, taking into account the considerable decrease of the diffusion length *L* in the film, it is easy to realize much better spatial resolution than in the case of a bulk material.



Figure 9. Probe energy dependences of the electron beam divergence *b* after passing Si films of 1 μ m (*1*) and 2 μ m (*2*) thick [21].

The observations of thin objects in a microcathodoluminescence mode may be fulfilled in a scanning transmission electron microscope. In this case the combination of appropriate findings with the possibility to visualize the crystal defects by the usual methods of a transmission electron microscopy permits one to study the peculiarities of recombination processes near individual defects of crystal structure, e.g., dislocations and stacking faults.

4. Instrumentation

A standard experimental setup intended to investigate a microcathodoluminescence includes a routine SEM or its transmission version (STEM), a system which collects light radiation and brakes it out the microscope column (collector system), a system for spectral analysis of radiation (usually monochromator), and a system which monitors a CL radiation. The collector system mostly consists of a mirror and lenses and is constructed to collect as many of emitted photons as possible. Photomultipliers are usually used to detect visual light, they comprise very efficient photodetectors with gain factors up to 10⁹. In the near IR range solid-state photodetectors are employed: Ge and Si photodiodes [22], and PbS and InSb photoresistors [23].

The regular and most widespread operating modes of such a setup consist in recording CL images of the objects and monitoring CL emission spectra. In the first case either total (integral or panchromatic) CL signal, which comes to the photodetector bypassing the monochromator, or a fixed spectral wavelength (monochromatic) CL signal, passing an appropriately tuned monochromator, are recorded. Here, the electron beam is scanned over the object surface. In the second case the electron beam is immovable, being positioned at a definite, chosen point on the object, and now the monochromator is scanned over the fixed spectral range. This procedure allows for getting a CL radiation spectrum emitted from a given point.

5. Information obtained from a cathodoluminescence signal

The CL characteristics which carry useful information about an object involve the total (integral) intensity of the CL radiation, CL spectra, a dynamic of the onset and decay of the CL radiation, and a polarization of the CL radiation. Let us consider the examples of information which may be obtained when recording one or another CL characteristic.

5.1 An integral cathodoluminescence intensity

This CL characteristic seems to be easiest for detecting. The image in this case represents nothing but a map of distribution over the object surface for a spatially dependent photonemission rate which may vary over the surface due to a variety of factors.

In accordance with (7) the intensity I_{CL} of CL radiation depends on four factors: A, B, η , and Δn , each being able to have nonuniform spatial distribution over the surface of the object. In general, the excess-carrier generation rate Δn has uniform distribution except in objects that have portions of the surface covered with opaque layers of different composition, e.g., metal contacts. This leads to the change of probe electron energy losses directly in the object material itself, and, accordingly, to the change of Δn .

The other three quantities are subjected to surface variations to a greater extent. Thus, the presence of contamination films or thickness variations of dielectric (cover) layers may result in changes in reflectivity of light at the object-vacuum interface, i.e. the value of *B*. In objects such as multilayer heterostructures, the local variations of light absorption coefficient may take place thereby affecting the *A* quantity. There is a variety of factors that may influence the local values of radiative $\tau_{\rm rr}$ and nonradiative $\tau_{\rm nr}$ recombination lifetimes thus producing the spatial variations of the internal quantum efficiency $\eta(r)$ over the object surface.

As a result, a 'routine' CL-image contrast may be provoked by an action of several of the above causes, so it is very important to separate the contrast due to geometrical effects from that due to irregularities in the luminescent properties (internal quantum efficiency). If thin objects are under investigation, the variations in the I_{CL} may result from changes in an object thickness leading to the alteration of a degree of the probe electron energy losses in the object.

The factors affecting the internal quantum efficiency, and information about which may be obtained from the I_{CL} signal, are the following:

(1) The dopant type and concentration

Such a dependence of the CL intensity measured from GaAs: Te on the tellurium concentration is shown in Fig. 10



Figure 10. Dependence of the CL intensity at 300 K on donor concentration for *n*-GaAs [24].

[24]. It is seen that the CL intensity is very sensitive to the tellurium concentration change that causes an appearance of contrast in CL images produced from the object regions with different dopant concentration. Another example is demonstrated in Fig. 11. The object was a polycrystalline PbS film which began to take its luminescent and photosensitive properties after annealing in air. The oxygen atoms penetrated deep into the film material inhomogeneously that is seen from the oxygen distribution map (Fig. 11a) obtained with an Auger microprobe [25], and this gave rise also to the inhomogeneity of the luminescent properties.



Figure 11. SEM images of a PbS polycrystalline film: (a) in Auger electrons whose energy corresponds to the oxygen Auger peak (the map of oxygen distribution over the film surface), and (b) in a panchromatic CL signal [25]. The horizontal size equals $32 \mu m$.

(2) The presence of structure defects

As an example, consider dislocations which most often reveal themselves as centres with higher nonradiative recombination rate, and assume that an additional nonradiative phenomenon is associated with a dislocation being characterized by a lifetime τ_d . Then the effective lifetime at a dislocation τ' , unlike the bulk lifetime τ_{bulk} , will be defined by the relationship

$$\frac{1}{\tau'} = \frac{1}{\tau_{\text{bulk}}} + \frac{1}{\tau_{\text{d}}}$$

and a dislocation contrast can be expressed as $C = 1 - \tau'/\tau_{bulk}$. Since $\tau' < \tau_{bulk}$, the dislocations appear in CL images in the form of black dots that has been shown for the first time in [26]. In highly doped semiconductors, another type of dislocation contrast, so-called 'dot and halo' contrast, was observed [3]. The possible causes of its appearance are connected either with dopant atoms segregation at dislocations surroundings due to decrease of dopant concentration (the right portion of the curve in Fig. 10), or with dislocation gettering effect that results in reducing the local concentration of nonradiative centres and leading to local CL enhancement due to decrease of images are presented in Fig. 12.



Figure 12. The integral CL images of dislocations in GaAs:Te: (a) as dark dots (the tellurium concentration is 10^{17} cm⁻³), and as dots and halos (the tellurium concentration is 10^{18} cm⁻³) [3]. The horizontal size runs to 41 µm (left) and 85 µm (right).

5.2 Cathodoluminescence spectra

The CL spectrum contains information about radiative recombination mechanisms and energy states responsible for the CL emission from a given object region. Information obtained from spectra strongly depends on temperature: on the one hand, the broadening of spectral lines at cryogenic temperatures is small, and spectral lines corresponding to the concrete recombination mechanisms are well resolved, but, on the other hand, a change of recombination mechanisms with temperature is also possible. At cryogenic temperatures an essential reduction in the rate of an electron-beam-induced damage of the object takes place. This fact goes especially important when such easy degradable materials like CdTe and HgCdTe are investigated.

With the availability of several radiative recombination mechanisms to reveal themselves in the object it is necessary to take into account that all of them compete with each other as well as with nonradiative recombination channel. This fact must be allowed for in cases when the excitation density changes in the experiment. At the expense of this the essential redistribution of recombination rates is possible between the separate recombination channels leading on occasion to considerable changes in the spectrum. This effect has to be taken into consideration when comparing CL spectra obtained from different regions of the object. The difference in nonradiative recombination rates may be a cause of the difference in the effective excitation density in these regions even at the invariable experimental conditions, and the result is an apparent change in the efficiency of the concrete recombination channel.

Selecting with a monochromator one or another part of the spectrum which corresponds to a particular radiative recombination mechanism, and using this signal for imaging, the distribution of this recombination mechanism over the object surface can be mapped. Such a picture may represent by itself a map of the separate phase distribution over the object surface, if the CL emission in a particular part of the spectrum corresponds to one phase at a time. This mode, if the phases are disposed (hidden) at different depths, may be used for nondestructive control of complex structures. It was demonstrated in [27], where the defects in the active layer of a InGaAsP double heterostructure were observed under the emitter layer (naturally, the emitter layer was thin enough in order that beam electrons could reach the active layer).

Advances in epitaxial growth techniques have led to a production of special (highly perfect) heterostructures such as quantum wells (QWs) and superlattices. Alternation of semiconducting layers with different band gaps, where it is possible to set conditions for harnessing quantum size effects, takes place in these structures, and a new type of intrinsic emission is possible there. A quantum well is a thin layer of a narrow-band-gap semiconductor placed between wide-bandgap semiconducting layers. A periodic array of such alternating layers with different band gaps constitutes a macroscopic quantum superlattice. In the first approximation the electron energy states in the QW can be found by simplest quantummechanical calculations, so that for a 'particle in the onedimensional box' and for infinitely deep well they are as follows:

$$E_n = \frac{\hbar^2}{2m} \left(\frac{\pi n}{L}\right)^2,\tag{10}$$

where L is the width of the narrow-band-gap layer. Of course, the potential well is not infinitely deep, so the wave functions die away exponentially into the potential barriers, and this may account for tunnelling through the barriers in a periodic structure of Esaki superlattice [28].

The important luminescent properties of such wells are pertinent to (1) strong dependence of the energy of the photon emitted by the transition from the electron states (Eqn 10) to the hole states on the well layer width L, and (2) since the electrons and holes are confined in the same well layer, the strongly bound excitons are formed there, and since the excitons tend to recombine radiatively, a radiative recombination efficiency in the well layer is high even up to room temperature. How high the sensitivity to the well-layer thickness alteration is, the following fact shows: changing the thickness only by one monolayer height leads to a variation in the exciton peak position by several nanometers what is easy to be measured. Therefore, it was proposed in Ref. [29] to use the quantum well exciton as a beam for monitoring the QW interface roughness. This is one of the most important problems connected with the growth of QW and superlattice structures because these structures consist wholly or in part of interfaces, and radiative and nonradiative recombination take place in the closest vicinity to the heterointerfaces. In such a manner the giant monolayergrowth islands were found at QW interfaces by using for



Figure 13. Illustration of the principle of CL imaging of an island-nature structure of single QW interfaces with monoatomic height islands on each interface (above); the CL spectrum with three peaks corresponding to areas of the QW layer different in thickness and a relief of the lower half of the interface are also shown below [29].

By variation of the accelerating voltage in the CL mode and observing the CL spectra it is possible to realize the unique feature of the CL technique — to get information about the object in the third dimension, i.e. along the coordinate directed in depth of the object (the so-called 3D microscopy) and not photholuminescence. This possibility is merely possessed by CL.

By varying the accelerating voltage a penetration depth of beam electrons as well as a distribution of their energy losses over the object (i.e. the electron-hole pair density distribution) change that alters the depth of the CL signal generation and contribution of the separate object layers to the total CL signal. This technique of a 'depth probing' is nondestructive and irreplaceable in studying ion-implanted semiconducting materials, since the penetration depths of electrons with energies between 1 and 20 keV are comparable to that of ions with energies from 50 to 300 keV usually used for technological purposes in semiconductor industry. This fact permits the use of the CL technique for investigating the implanted dopant and induced lattice-damage defect distribution in depth.

The examples in Fig. 14 [30] demonstrate the transformation of the CL spectra of a 100 keV-Xe⁺-implanted CdTe:In sample with variation of the SEM accelerating voltage. The objects were etched in depth to about 140 nm. Defects induced by implantation cause the appearance of a 1.2 eV luminescence band observed from depths by several hundred nanometers exceeding the implant-ion range projected to the surface normal (the so-called secondary effect of 'driving in', i.e. an additional moving forward of implanted atoms in the object depth through the influence of shooting ions). The appearance of the implantation-induced band is accompanied by quenching of a 1.4 eV luminescence which is believed to be due to transitions involving defect-impurity complexes in CdTe. This may be caused by two factors: (1) a non-



Figure 14. CL spectra (at 80 K) of 100 keV-Xe⁺- implanted and etched in depth of 140 nm CdTe:Im sample as a function of the electron beam energy: (a) -20 keV, (b) -7.368 keV, (c) -2.714 keV, and (d) -1.0 keV. The electron penetration depth is, accordingly: 1.44, 0.25, 0.044, and 0.0078 μ m [30].

radiative recombination efficiency growth at the expense of implantation-induced structure defects, what is usually 'cured' by the further annealing, and (2) the competition of a new (1.2 eV) radiative recombination channel.

The possibilities to obtain CL spectra at low temperatures broaden our horizons as regards the analysis of the separate object parts. Actually, with the exception of emission bands which correspond to interband transitions, free exciton recombination as well as recombination through the localized states in the forbidden band, formed by dislocations, all the other emission bands result from recombination with the participation of impurity atoms. For each semiconducting material the levels in the gap due to different impurities can be tabulated. If transitions via these levels are radiative, then the presence of the corresponding lines in the emission spectra will allow for conclusions about the presence of a given impurity in the analyzed volume of the object to be made, i.e. to perform the qualitative microanalysis similar to X-ray microanalysis. The sensitivity of such a microanalysis is of several orders in magnitude higher than that of X-ray microanalysis (the detectability limit runs to 10¹⁴ cm⁻³ or less [31]). If the colour CL technique is used (see, for example, [32]), this limit may be further lowered due to the higher sensitivity of a human eye to colour variations.

As far as a quantitative CL analysis is concerned, there are no practically appropriate theoretical elaborations for it unlike the situation in an X-ray microanalysis, where the quantitative methods were well worked out. Dependences connecting impurity concentrations, CL emission intensities, and band structure parameters are yet unknown. One of the essential difficulty is involved in accounting for a nonradiative recombination, and in controlling the effective excitation density. A possible way to meet above difficulty is to compare the intensity of the line considered with another reference-line intensity which is determined by the carrier lifetime. In the case of linear recombination, the intrinsic luminescence line or, if it is impossible, a group of lines or a band-edge emission may be used for a reference line, as it has been suggested in Ref. [33]. The intrinsic luminescence line can be used in some cases for determination of a matrix composition too. One of the advantages in using the local CL spectra for this purpose as compared to X-ray ones is the possibility to reach a better spatial resolution at the sacrifice of operation with lower accelerating voltages.

It should be noted that there exists one more unique possibility inherent in the CL technique which permits us to investigate kinetics of a luminescence onset and decay, and

which unlike the photoluminescence case allows the change of the excitation pulse length without altering the rise and decay times of the pulse to be made. The knowledge of the lifetimes of nonequilibrium carriers in different states is very important, both from the fundamental and applied points of view. The properties and performance limits of many devices of solid-state electronics depend on capture cross-sections. It is very difficult to calculate them because of a variety of parameters. A time-resolved local CL combines the advantage of an electron probe instrument to study small volumes in the object with the possibility to gain information on capture processes by observing the time-dependent behaviour of separate lines and bands in CL spectra. The realization of this possibility has been come to in Ref. [34]. Experiments have been conducted on n- and p-GaAs, whereas CL spectra and transient relaxation processes have been investigated depending on the excitation pulse length τ_p . The CL spectra taken just after excitation with 2, 50 and 500 ns pulses (Fig. 15) differed essentially, and the number of working recombination mechanisms was larger after the longer excitation pulses. Following the 2 ns-pulse, the CL spectrum from the *n*-type object (Fig. 15a) did not show any features connected with the existence of neutral acceptors, whereas the spectrum from the *p*-type object (Fig. 15b) — neutral donors. Thus, up to 2 ns after excitation onset, recombination processes proceeded on bound minority carriers are completely absent, although free minority carriers are present. The capture of free minority carriers into bound states occurs at a much later time. These spectra directly visualized for the first time the temporal evolution of the capture processes.

The transients of the CL signal after the end of the exciting pulse, naturally, depended strongly on τ_p as long as a quasisteady state was not yet reached ($\tau_p \sim 1 \mu s$). At $\tau_p = 2 ns$, the integral CL intensity decayed with a time constant of about 1 ns (Fig. 16) corresponding to the relaxation time of the exciton bound on a donor (see also Fig. 15a). With the τ_p



Figure 15. Time-resolved local CL spectra of lightly doped *n*-GaAs:Sn (a) and high-purity *p*-GaAs (b) at 5.3 K. The spectra were taken after the time intervals of 2 (I), 50 (2) and 500 (3) ns from the beginning of electron excitation pulses. A⁰ is a neutral acceptor, and D⁰ is a neutral donor [34].



Figure 16. Transients of the panchromatic CL intensity of lightly doped *n*-GaAs:Sn sample after excitation with electron pulses of length 2(1), 50(2), and 500 ns(3) at 5.3 K [34].

growth, the relaxation processes at neutral acceptors became predominant.

5.3 Polarization of cathodoluminescence radiation

The CL polarization has been observed in a few works. For the first time, it was apparently revealed on $Zn_2SiO_4(Mn)$ [35]. It was found that radiation from the grains with the *c*axis parallel to the object surface was highly polarized, and the electric vector **E** was perpendicular to the *c*-axis. If the *c*axis of a crystal was perpendicular to the surface, the CL radiation from it was unpolarized. An interesting detail has been pointed out, but without an explanation — the peaks in the spectra from the common crystal differed in position by 10 nm, if spectra were taken with the analyzer arrangements distinguished by 90°.

In diamonds of type IIb, the blue CL radiation from linear defects oriented in the $\langle 110 \rangle$ direction was observed, and it was strongly polarized with the electric vector E being always in the above direction [36]. Later on, the same authors identified these defects as dislocations [37]. It is interesting that nearly all the CL emission from such diamonds originated from dislocations, although not all of them were luminescent [38]. The CL emission from individual dislocations in a type IIb diamond was observed at the wavelength of 435 nm, had maximum intensity at 170 K, and was polarized. All the luminescent dislocations were dissociated, and dislocations directed along (110) featured a strong luminescence. The vector E was parallel to the dislocation, regardless of the Burgers-vector direction. A donor-acceptor pair recombination model was suggested, taking into account the different distances between pairs and the dislocation core [39].

A polarized local CL has also been observed on several minerals [40]. For sphalerite, the degree of polarization reached 100%. In the work cited, the two images of minerals were obtained simultaneously with different (by 90°) analyzer positions before CRT in each channel.

In more later works (e.g., see Ref. [41]) the polarization of the local CL has been used to determine the spatial distribution of the stress tensor in GaAs/Si structures near and far from microcracks. The structures exhibited uniaxial and biaxial strains, and strain-induced splitting of the GaAs valence band into a heavy-hole (hh) and a light-hole (lh) subbands occurred, as it was shown by studying the hh- and lh-exciton cathodoluminescence.

6. Perspectives of the technique development and conclusions

For reasons of space we could not touch in this review all the aspects of a CL microscopy and the CL phenomenon as a whole. The problem of artifacts that may arise in CL experiments was not discussed as well. Also, we omitted the treatment of the question of CL technique application to a geology and biology, so far as their objects need some specific preparation just before installation into the microscope.

As far as the nearest perspectives of a CL microscopy are concerned, the following should be, apparently, mentioned here:

1. Transition from qualitative methods to quantitative ones should be detailed. It concerns the determination of both the values of separate electrophysical parameters of the object, and luminescent dopants concentration. Taking into account that the sensitivity of the CL analysis many times (at least 10^5) exceeds that of an X-ray microanalysis, the CL technique would permit one to control the real dopant concentration in some semiconducting materials.

2. Extension of the working wavelength range of the CL investigations into the near ultraviolet and infrared parts of the spectrum beyond $\lambda = 5 \ \mu m$. In the latter case special types of photodetectors with a large working area and higher sensitivity have to be constructed.

Acknowledgment. The author is very much indebted to Prof. V S Vavilov for his useful and constructive remarks in refereeing this article.

References

- 1. Moskvin A V *Katodoluminestsentsiya* Chaps 1, 2 (Cathodoluminescence) (Moscow-Leningrad: Gostekhizdat, 1948–1949),
- Pankove J I Optical Processes in Semiconductors (Englewood Cliffs, New Jersey: Prentice-Hall, 1971) [Translation into Russian; Moscow: Mir, 1973]
- Yacobi B G, Holt D B Cathodoluminescence Microscopy of Inorganic Solids (New York, London: Plenum Press, 1990)
- a) Gareeva A R et al. *Izv. Ros. Akad. Nauk, Ser. Fiz.* **59** (9) 200 (1995) [*Bull. Rus. Acad. Sci., Ser. Phys.* **59** (9) (1995)]; b) Mitdank R et al. *Semicond. Sci. Technol.* **10** 835 (1995)
- 5. Bethe H Ann. Phys. (Leipzig) **5** 325 (1930)
- 6. Kanaya K, Okayama S J. Phys. D: Appl. Phys. 5 43 (1972)
- Saparin G V et al. Vestnik Mosk. Univ., Fiz. Astron. 24 (3) 56 (1983) [Moscow Univ. Phys. Bull. 24 66 (1983)]
- 8. Yacobi B G, Holt D B J. Appl. Phys. 59 R1 (1986)
- 9. Bresse J-F Mater. Sci. Eng. B 24 229 (1994)
- Gareeva A R et al. Vestnik Mosk. Univ., Fiz. Astron. 31 (2) 33 (1990) [Moscow Univ. Phys. Bull. 31 (2) (1990)]
- El'tekov V A, Gvozdover R S, Petrov V I *Izv. Akad. Nauk SSSR*, Ser. Fiz. **52** (7) 1376 (1988) [Bull. Acad. Sci. USSR, Phys. Ser. **52** (7) 1376 (1988)]
- Gareeva A R, Petrov V I, Chizhov G A Vestnik Mosk. Univ., Fiz. Astron. 30 (1) 23 (1989) [Moscow University Physics Bull. 30 (1) (1989)]
- Leisegang S *Elektronenmikroskopie* (Berlin-Göttingen-Heidelberg: Springer-Verlag, 1956) [Translation into Russian (Moscow: Izd. In. Liter., 1960) p. 136]
- Glaser W Grundlager der Elektronenoptik (Wien: Springer-Verlag, 1952) [Translation into Russian (Moscow: Gostekhteorizdat, 1957) p. 703]
- Spivak G V et al. Pis'ma Zh. Eksp. Teor. Fiz. 21(1) 38 (1975) [JETP Lett. 21 (1) 17 (1975)]
- 16. Araujo D et al. Mater. Sci. Eng. B 24 124 (1994)
- 17. Warwick C A J. de Phys. IV 1 C6-117 (1991)
- Spivak G V et al. Fiz. Tekh. Poluprovodn. 14 1934 (1980) [Sov. Phys. Semicond. 14 (10) 1151 (1980)]
- 19. Wada K et al. Jpn. J. Appl. Phys. 127 L1952 (1988)

- 20. Samuelson L et al., J. Vac. Sci. Technol. B 12 2521 (1994)
- Goldstein J I et al. in Proc. of the 10th Annual Scanning Electron Microscopy Symposium, Chicago, IL., USA, 1977 (Chicago, IL.: IIT Res. Inst., 1977) p. 315
- 22. Koschek G, Kubalek E Scanning 7 (4) 199 (1985)
- Petrov V I et al. Izv. Akad. Nauk. SSSR, Ser. Fiz. 51 (3) 443 (1987) [Bull. Acad. Sci. USSR, Phys. Ser. 51(3) 25 (1987)]
- 24. Cusano D A Sol. State Commun. 2 353 (1964)
- Petrov V I, Prokhorov V A *Izv. Akad. Nauk SSSR, Ser. Fiz.* **48** (2)
 263 (1984) [*Bull. Acad. Sci. USSR, Phys. Ser.* **48** (2) 46 (1984)];
 Petrov V I, Prokhorov V A, Yunovich A E *Fiz. Tekh. Poluprovodn.* **18** (3) 484 (1984) [*Sov. Phys. Semicond.* **18** (3) 300 (1984)]
- Kyser D F,Wittry D B, in *The Electron Microprobe* (Eds T D McKinley, K F J Heinrich, D B Wittry) (New York: Wiley, 1964) p. 691.
- Petrov V I et al. Izv. Akad. Nauk SSSR, Ser. Fiz. 48 (12) 2404 (1984)
 [Bull. Acad. Sci. USSR, Phys. Ser. 48 (12) 107 (1984)]
- 28. Esaki L, Tsu R IBM J. Res. Develop. 14 61 (1970)
- 29. Bimberg D et al. Superlattices and Microstructures 4 257 (1988)
- 30. Norris C B, Barnes C E, Zanio K R J. Appl. Phys. 48 1659 (1977)
- 31. Holt D B, Saba F M Scanning Electron Microscopy (USA) Pt. 3 1023 (1985)
- 32. Spivak G V, Saparin G V, Antoshin M K Usp. Fiz. Nauk 113 695 (1974) [Sov. Phys. Usp. 17 595 (1975)]
- Gippius A A, Vavilov V S, in *Ionnaya implantatsiya v poluprovodnikakh i drugikh materialakh* (Ion Implantation in Semiconductors and Other Materials) (Vilnus: Izd. Minvuz Lit. SSR, 1985) p. 59
- 34. Bimberg D et al. Phys. Rev. B 31 7788 (1985)
- 35. Bhalla R J R S, White E W J. Appl. Phys. 41 2267 (1970)
- 36. Kiflawi I, Lang A R Phil. Mag. 30 219 (1974)
- 37. Kiflawi I, Lang A R Phil. Mag. 33 697 (1976)
- 38. Pennycook S J et al. Phil. Mag. A 41 589 (1980)
- 39. Yamamoto N et al. Phil. Mag. B 49 609 (1984)
- 40. Spivak G V et al. Izv. Akad. Nauk SSSR, Ser. Fiz. **41** 876 (1977) [Bull. Acad. Sci. USSR, Phys. Ser. (1977)]
- 41. Rich D H et al. Phys. Rev. B 43 6836 (1991)